



**T.C. İSTANBUL TİCARET
ÜNİVERSİTESİ**

FEN BİLİMLERİ ENSTİTÜSÜ

**TFRS 9 VE TEMERRÜT OLASILIĞI MODELLEMESİ
BANKACILIK ALANINDA KARŞILAŞTIRMALI
BİR UYGULAMA**

Celil TAŞKIN

**Danışman
Prof. Dr. Münevver TURANLI**


**YÜKSEK LİSANS TEZİ
İSTATİSTİK ANABİLİM DALI
İSTANBUL - 2019**

KABUL VE ONAY SAYFASI

Celil TAŞKIN tarafından hazırlanan "**TFRS 9 ve Temerrüt Olasılığı Modellemesi Bankacılık Alanında Karşılaştırmalı Bir Uygulama**" adlı tez çalışması 09/07/2019 tarihinde aşağıdaki jüri üyeleri önünde başarı ile savunularak, İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü **İstatistik Anabilim Dalı**'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman

Prof. Dr. Münevver TURANLI
İstanbul Ticaret Üniversitesi



Jüri Üyesi

Prof. Dr. Serpil Ergün BÜLBÜL
Marmara Üniversitesi



Jüri Üyesi

Prof. Dr. Ünal Halit ÖZDEN
İstanbul Ticaret Üniversitesi



Onay Tarihi: 02.09.2019



Doç. Dr. Serkan ÇANKAYA
Enstitü Müdürü Vekili

AKADEMİK VE ETİK KURALLARA UYGUNLUK BEYANI

İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

02.09.2019

Celil TAŞKIN



İÇİNDEKİLER

	Sayfa
İÇİNDEKİLER	iv
ÖZET	vi
ABSTRACT	vii
TEŞEKKÜR	viii
ŞEKİLLER	ix
TABLolar	x
SİMGELEr VE KISALTMALAR	xi
1. GİRİŞ	1
2. LİTERATÜR ÖZETİ	4
3. TFRS 9 VE TEMERRÜT OLASILIĞI MODELLEMESİNDEKİ KAVRAMLAR	12
3.1 Beklenen Kredi Zararı	12
3.2 Finansal Araçlar Raporlama Standardı	13
3.3 Kredi Riski	15
3.4 Temerrüt Olasılığı	16
4. TFRS 9 VE TEMERRÜT OLASILIĞI MODELLEMESİNDE KULLANILAN İSTATİSTİKSEL YÖNTEMLER	18
4.1 Lojistik Regresyon Analizi	18
4.1.1 Lojistik regresyon analizinin varsayımları	20
4.1.2 Lojistik regresyon analizinin avantajları ve kısıtları	26
4.2 Cox Regresyon Analizi	27
4.2.1 Cox regresyon analizinin varsayımları	29
4.2.2 Cox regresyon analizinin avantajları ve kısıtları	30
5. UYGULAMA	32
5.1 Uygulamannın Amacı ve Hipotezi	32
5.2 Uygulamadaki Modelleme Yaklaşımı	32
5.2.1 Modellerde kullanılan temerrüt tanımı	33
5.2.2 Modellerin kurulacağı dönem	33
5.2.3 Modellerde yer alan müşteriler	33
5.2.4 Modellere ilişkin örneklemin oluşturulması	36
5.2.5 Modellerdeki değişkenler	37
5.2.5.1 Bağımlı değişken	38
5.2.5.2 Bağımsız değişkenler	38
5.2.6 Modellerde kullanılan verilerin analizi	39
5.2.7 Modelleri değerlendirme analizleri	40
5.2.7.1 Roc eğrisi ve gini katsayısı analizi	40
5.2.7.2 Hata matrisi analizi	41
5.3 Lojistik Regresyon Analizi Uygulaması	42
5.3.1 Değişken analizleri	42
5.3.2 Etkileşimli gruplama	43
5.3.3 Öğrenme ve doğrulama veri setleri	47
5.3.4 Değişken seçimi ve regresyon	48
5.3.5 Modelin performansı ve validasyonu	52
5.4 Cox Regresyon Analizi Uygulaması	55
5.4.1 Değişken analizleri	60
5.4.2 Etkileşimli gruplama	61
5.4.3 Öğrenme ve doğrulama veri setleri	62
5.4.4 Değişken seçimi ve regresyon	63

5.4.5 Modelin performansı ve validasyonu	65
6. SONUÇ VE ÖNERİLER.....	68
KAYNAKLAR	70
EKLER.....	74
Ek A. Değişkenler Listesi.....	74
Ek B. Lojistik Regresyon Analizi SAS Çıktıları.....	75
Ek C. Cox Regresyon Analizi SAS Çıktıları.....	87
ÖZGEÇMİŞ	102



ÖZET

Yüksek Lisans Tezi

TFRS 9 VE TEMERRÜT OLASILIĞI MODELLEMESİ BANKACILIK ALANINDA KARŞILAŞTIRMALI BİR UYGULAMA

Celil TAŞKIN

İstanbul Ticaret Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Prof. Dr. Münevver TURANLI

2019, 113 sayfa

Bu çalışmada, 2018'de yürürlüğe giren Finansal Araçlar Raporlama Standardının bankacılığa etkisi temerrüt olasılığı bağlamında ele alınmıştır. Bu Standart gereği, bankaların beklenen kredi zararı hesaplamasında kullanılmak üzere raporlama döneminden sonraki on iki ayda ve kredi ömrü boyunca müşteriler için temerrüt etme olasılıklarının hesaplanması gerekmektedir. Temerrüt olasılığı modellemesinde lojistik regresyon analizi yaygın olarak kullanılan bir tekniktir. Gerçek bir veri seti ile lojistik regresyon analizi uygulanarak temerrüt olasılığı hesaplama süreci aşama aşama ele alınmıştır. Bu kapsamda veri seti, modelleme yaklaşımı ve lojistik regresyon analizi hakkında bilgi verildikten sonra uygulama anlatılmıştır. Cox regresyon analizi Türkiye'de bankacılık alanında çok kullanılsa da lojistik regresyon analizine ciddi bir alternatiftir ve bazı açılardan daha kullanışlıdır. Bu yüzden Cox regresyon analizi anlatıldıktan sonra uygulamasına yer verilmiş ve sonuç bölümünde bu iki analizin çıktıları karşılaştırılmıştır.

Anahtar Kelimeler: Cox regresyon, kredi riski, lojistik regresyon, temerrüt olasılığı, TFRS 9.

ABSTRACT

M.Sc. Thesis

IFRS 9 AND PROBABILITY OF DEFAULT MODELLING A COMPARATIVE STUDY IN THE BANKING SECTOR

Celil TAŞKIN

**İstanbul Commerce University
Graduate School of Applied and Natural Sciences
Department of Statistics**

Supervisor: Prof. Dr. Münevver TURANLI

2019, 113 pages

In this study, the impact of the Financial Instruments Reporting Standard, which was put into effect in 2018, on banking was discussed in terms of the probability of default. In accordance with the standard, in order to use for calculating the expected credit loss of the banks, the probability of defaulting for customers is required to be calculated in the twelve months following the reporting period and over the lifetime of the loan. Logistic regression is a commonly used technique in probability modeling. The process of calculating the probability of default by using logistic regression with a real data set has been discussed step by step. In this context, after giving information about data set, modeling approach and logistic regression, application was explained. Although Cox regression is not commonly used in the banking sector in Turkey, it is an important alternative for logistic regression and more useful in some ways. Therefore, Cox regression was initially introduced and then application took place. In the conclusion, the results of these two techniques were compared.

Keywords: Cox regression, credit risk, IFRS 9, logistic regression, probability of default.

TEŐEKKÜR

Yüksek lisans ders döneminde ve tez hazırlığı safhasındaki ilgi ve desteęi için nezaket ve bilimin keřiștięi noktada duran iyi bir insan ve iyi bir akademisyen olan danıřman hocam Prof. Dr. Münevver TURANLI'ya teőekkür ederim.

Keyifli bir sohbet havasında kendilerinden ders aldığım deęerli hocalarım Prof. Dr. Ünal Halit ÖZDEN, Doç. Dr. Özlem Deniz BAŐAR ve Dr. Öğr. Üyesi Seda Baędatlı KALKAN'ı da Őükranla zikretmem gerekiyor.

Celil TAŐKIN
İstanbul, 2019



ŞEKİLLER

	Sayfa
Şekil 4.1. Cox regresyon analizi sansür çeşitleri	28
Şekil 5.1. Skor derecelendirme	55
Şekil 5.2. Grafiksels testler	58
Şekil 5.3. Validasyon sonuçları	67



TABLolar

	Sayfa
Tablo 5.1. Elemeler öncesi gözlem setinin dağılımı.....	36
Tablo 5.2. Elemeler sonrası gözlem setinin dağılımı.....	37
Tablo 5.3. WoE değerleri	44
Tablo 5.4. Lojistik modeldeki değişkenlerin gini katsayıları ve bilgi değerleri.....	46
Tablo 5.5. Gerçek veri seti.....	47
Tablo 5.6. Öğrenme ve doğrulama veri setleri.....	47
Tablo 5.7. Lojistik regresyon modelleme girdi değişkenleri.....	49
Tablo 5.8. Lojistik regresyon analizinde değişkenlerin modele giriş sırası.....	50
Tablo 5.9. Lojistik regresyon analizi parametre katsayıları	51
Tablo 5.10. Lojistik regresyon modeline dair temel istatistikler	52
Tablo 5.11. Kesim noktalarına göre hata matris tabloları	53
Tablo 5.12. Zaman etkileşimli değişkenlerin analizi	58
Tablo 5.13. Genişletilmiş cox regresyon modelindeki gözlemlerin dağılımı.....	59
Tablo 5.14. Genişletilmiş cox regresyon modelindeki değişkenlerin gini katsayıları ve bilgi değerleri.....	61
Tablo 5.15. Genişletilmiş cox regresyon öğrenme ve doğrulama veri setleri	62
Tablo 5.16. Genişletilmiş cox regresyondaki değişkenlerin modele giriş sırası	64
Tablo 5.17. Genişletilmiş cox regresyon analizi parametre katsayıları	65
Tablo 5.18. Genişletilmiş cox regresyon modeline dair temel istatistikler	66

SİMGELER VE KISALTMALAR

BDDK	Bankacılık Düzenleme ve Denetleme Kurumu
BKZ	Beklenen Kredi Zararı
IFRS 9	Uluslararası Finansal Raporlama Standardı 9
İDD	Kredi Riskine Esas Tutarın İçsel Derecelendirmeye Dayalı Yaklaşımlar ile Hesaplanmasına İlişkin Tebliğ
İMKB	İstanbul Menkul Kıymetler Borsası
KGK	Kamu Gözetimi Muhasebe ve Denetim Standartları Kurumu
MLR	Multinominal Lojistik Regresyon
TBB	Türkiye Bankalar Birliği
TFRS 9	Türkiye Finansal Raporlama Standardı 9
TMSF	Tasarruf Mevduatı Sigorta Fonu
TO	Temerrüt Olasılığı



1. GİRİŞ

Finans sektöründe müşterinin kredi değerliliğinin ölçülmesi için uzman görüşünden istatistik temelli çeşitli ölçümlere kadar birçok farklı yöntem mevcut olmakla birlikte genellikle finansal kurumların iş modelleri çerçevesinde melez modeller tercih edilmektedir. Böylece karar alınırken istatistikî modellerden - giderek artan bir oranla- istifade edilmekte ve nihai karar alıcı olarak insanî yargı devreye girmektedir.

Kredi değerliliği ölçülürken en yaygın kullanılan risk bileşeni, müşterinin sonraki bir yıl içinde yüzde kaç ihtimalle temerrüt edeceği. Müşteri bazlı hesaplanan temerrüt olasılıkları, derecelendirme notuna çevrilmekte ve karar alınırken bu derecelendirme notları kullanılmaktadır. Her bir derecelendirme notunun başlangıç ve bitiş temerrüt olasılıkları bir skala çerçevesinde belirlenmekte ve müşteri için hesaplanan temerrüt olasılığı skalada hangi aralığa denk geliyorsa ilgili derecelendirme notu müşteriye atanmaktadır. Böylece müşterileri, kredi değerliliklerine göre objektif bir şekilde sınıflandırmak mümkün olmaktadır.

Finansal kurumlar, temerrüt olasılığı modellerini giderek artan bir oranda daha etkin bir şekilde kullanmaktadırlar. Bunun belli başlı sebepleri şöyle özetlenebilir. Birincisi, karar alırken objektif olması gereken karar alıcılar, çeşitli nedenlerle sübjektif kararlar alabilirler. Örneğin, kredi verilmemesi gereken bir müşteriye hedef baskısından dolayı hedefi tutturabilmek için kredi tahsisi yapılabilir. İstatistikî modellerin kullanılmasıyla birlikte bu hatalı kararların önüne geçilmesi kolaylaşmaktadır. İkincisi, limit tahsis edilirken veya limit arttırımı veya azaltımı yapılırken daha objektif ve daha hızlı karar alınabilmektedir. Üçüncüsü, sermaye yeterliliği hesaplamalarında içsel derecelendirmeye dayalı kredi riski hesaplama yaklaşımına geçilmesi hususunda yasal izinlerin alınabilmesi için içsel derecelendirme modellerinin banka genelinde yaygın bir şekilde kullanılması gerekmektedir. Finansal kurumlar ya mevzuatta belirtilen standart yaklaşımla sermaye yeterliliği hesaplaması

yapacaklar ya da sermaye yeterliliği hesaplamasında en önemli bileşen olan kredi riskine esas tutar hesaplamasında kendi içsel derecelendirme notlarını baz alabileceklerdir. İkincisini tercih ettiklerinde sabit oranlara kıyasla daha avantajlı olacakları beklendiği için bankalar derecelendirme modelleri geliştirmeye ve yaygın kullanımına önem vermektedirler. Dördüncü ve en önemli neden, yeni bir yasal zorunluluktan kaynaklanmaktadır. İçsel olarak üretilen temerrüt olasılıkları, TFRS 9 (Türkiye Finansal Raporlama Standardı 9: Finansal Araçlar) kapsamında, beklenen kredi zararlarının hesaplanmasında 1 Ocak 2018'den itibaren kullanılmaya başlanmıştır. Bankacılık Düzenleme ve Denetleme Kurumu (BDDK)'dan izin alarak TFRS 9 uygulamayacak bazı bankalar olacak olsa da, çoğu banka, yasal zorunluluk olarak her bir müşterisi için temerrüt olasılığı hesaplamak zorundadır. Bu kapsamda bankaların TFRS 9'a uyumlu temerrüt olasılığı (TO) modelleri geliştirmeleri yasal bir zorunluluk haline gelmiştir. Araştırmanın yoğunlaşacağı temel problem, yasal zorunluluk nedeniyle ortaya çıkan bankaların TFRS 9 uyumlu TO model ihtiyacıdır.

Yukarıda bahsedilen nedenlerden dolayı Türkiye'deki bankaların TO modellemesi hususunda çok geriye gitmeyen bir geçmişi ve bir miktar bilgi birikimi mevcuttur. Türkiye'de TO modellemesinde en yaygın kullanılan yöntem, lojistik regresyon analizi yöntemidir. İlerleyen bölümlerde değinileceği üzere, lojistik regresyon analizinin bazı kısıtları olsa da yaygın kullanılmasının makul sebepleri bulunmaktadır. Türkiye'de henüz pek fazla gündem olmasa da İngilizce literatür araştırıldığında akademi ve finansal kuruluşlarda lojistik regresyon analizine alternatiflerin neler olacağı tartışıldığı görülmektedir. Söz konusu literatürde Cox regresyon analizi bir alternatif olarak önerilmektedir. Bu tez çalışmasıyla bu alternatif gündeme getirilerek karşılaştırmalı bir şekilde analiz edilecektir.

TFRS 9, Türk ekonomisinin yapı taşlarından olan bankaları önemli ölçüde etkileyecektir. Bankaların 2018 ilk çeyrek denetim raporlarında yayınladıkları TFRS 9 geçiş etkisi verileri bunu göstermektedir. Bu yüzden TFRS 9 ve onun doğurduğu istatistikî modelleme ihtiyacı üzerinde akademik çalışmaların

yapılması gerekmektedir. Muhasebe açısından TFRS 9'u ele alan öncül çalışmalar mevcut olsa da, bankaların modelleme ihtiyacı açısından TFRS 9'u ele alan bir çalışma -görebildiğimiz kadarıyla- literatürde yoktur. Tez, bu açıdan ilk çalışmalardan birisi olacaktır.

Tezi ana hatlarıyla özetlemek gerekirse, ikinci bölümde literatür taraması yapılarak konu ile ilgili tezlere ve karşılaştırmalı makalelere yer verilmiştir. Üçüncü bölümde ilgili kavramlar üzerinde durularak kavramların içerikleri netleştirilmiştir. Dördüncü bölümde araştırma yöntemi ve modelleme yaklaşımı açıklandıktan sonra, varsayımlar üzerinde durulmuştur. Uygulanacak olan istatistiksel analizler anlatılmış ve lojistik regresyon analizi ile Cox regresyon analizi detaylıca ele alınmıştır. Beşinci bölümde lojistik regresyon analizi ve Cox regresyon analizi uygulamaları yer almaktadır. Sonuç bölümünde ise uygulama sonuçları özetlenmiş ve ardından sonraki çalışmalara ve modelleme ile uğraşanlara yönelik önerilerde bulunulmuştur.

2. LİTERATÜR ÖZETİ

Çalışmada kullanılacak olan Cox regresyon analizi daha çok Biyoistatistik çalışmalarında kullanılmakta ve bu çalışmalarla ilgili geniş bir literatür bulunmaktadır. Cox regresyonun finans alanında kullanılması Türkiye için nispeten yeni bir olgudur. Bu yüzden çok fazla yayın yoktur. Tezin kapsamı doğrultusunda burada sadece finansla ilgili çalışmalara yer verilecektir. Finans alanında Cox regresyon analizi ve/veya lojistik regresyon analizinin kullanıldığı belli başlı Türkçe tez çalışmalarını şöyle özetleyebiliriz:

Bülbül (1999), diskriminant analizi ve lojistik regresyon analizi gibi çok değişkenli modeller ile Cox regresyonu analizini karşılaştırıp, Cox regresyon analizinin onlara bir alternatif olup olamayacağı üzerinde durmuştur. Bu çalışmada, iflas veya üst üste 3 yıl zarar durumlarından birisi gerçekleştiğinde, işletme başarısız olarak sınıflandırılmıştır. İstanbul Menkul Kıymetler Borsası (İMKB)'na kayıtlı olan ve finansal kurum (banka veya sigorta şirketi) olmayan işletmeler üzerinden analizler yapılmış, 24 tane finansal oran ve şirketin yaş değişkenleri üzerinden şirketlerin başarısızlık zamanları tahmin edilmeye çalışılmıştır. Çalışma sonucunda, Cox regresyon analizinin finansal başarısızlık analizlerinde, diğer çok değişkenli analiz yöntemlerine alternatif olarak kullanılabileceği sonucuna varılmıştır.

Işık (2007), Türkiye'deki ticari bankaların finansal başarısızlıklarını etkileyen faktörleri Cox regresyon kullanarak analiz etmiştir. İlgili tezde finansal başarısızlığın tanımlanmasında, bankanın Tasarruf Mevduatı Sigorta Fonu (TMSF)'na devri, bankacılık izninin iptal edilmesi, başka bir bankaya devredilmesi, çoğunluk hisselerinin satışı ve neticesinde unvan değişikliği gibi kriterler kullanılmıştır. Sermaye yeterliği ile ilgili 4 rasyo, aktif kalitesi ile ilgili 6 rasyo, likidite yeterliliği ile ilgili 2 rasyo, karlılıkla ilgili 3 rasyo, gelir-gider yapısıyla ilgili 3 rasyo ve banka yaş değişkeni olmak üzere toplam 19 değişken ile modelleme çalışması yapılmıştır. Çeşitli alternatif modeller geliştirilerek

karşılaştırmalar yapılmış ve toplam kredilerin toplam aktiflere bölünmesiyle oluşturulan değişkenin en anlamlı olduğu bir modelin en uygun olduğuna karar verilmiştir.

Börüban (2009), çalışmasında lojistik regresyon analizi ve diskriminant analizi kullanarak firmaların mali başarısızlıklarını öngören modeller kurmuştur. Literatürde çok farklı mali başarısızlık tanımları olduğuna vurgu yapıldıktan sonra ilgili çalışmada mali başarısızlık kavramı, “Firmanın kredi borçlusu olduğu bankaya olan yükümlülüklerini yerine getirememesi sonucu yakın izlemeye alınması” olarak tanımlanmıştır. Diğer taraftan çalışmada analiz oldukça küçük bir örneklem (80 başarılı firma ve 80 başarısız firma) ile yapılmıştır. Firmaların finansal verileri üzerinden 8 tane finansal oran hesaplanmış ve bu 8 finansal oran bağımsız değişken olarak kullanılmıştır. Borç ödeme gücü ile ilgili 2 rasyo, mali yapı ile ilgili 2 rasyo, karlılıkla ilgili 2 rasyo ve aktivite oranlarıyla ilgili 2 rasyo kullanılmıştır. İki farklı analiz türü (lojistik regresyon analizi ve diskriminant analizi) ile geliştirilen modellerin öngörü gücü birbirlerine yakın bulunmuştur.

Tezergil (2009), sigorta sektörüne ait verilere çok aşamalı lojistik regresyon analizi uygulayarak kişilerin bir ürünü (bireysel emeklilik sigortası) satın alma kararlarını etkileyen değişkenleri analiz etmiştir. Tahmin edilmeye çalışılan durum, bireylerin yeni ürünü satın alıp almayacağı ve bu karar üzerinde müşteri temsilcilerinin etkisinin olup olmadığıdır. Bu kapsamda literatür taraması ve alanın uzmanlarıyla yapılan görüşmeler neticesinde bir anket formu oluşturulmuş ve anket uygulanmıştır. Eksiklikler olan anketler kapsam dışında tutulmuş ve 761 tane müşteri anketi ile 79 tane müşteri temsilcisi anketi üzerinden veri seti oluşturulmuştur. Kurulan modeller sonucunda müşterilerin satın alma kararlarında müşteri temsilcilerinin etkili olduğu görülmüştür. Müşteri özellikleri açısından kişinin çevresindekilerin ilgili ürünü satın almış olması, ilgili ürün hakkında bilgi sahibi olma ve gelir düzeyi ile ilgili ürünü satın alma ihtimali arasında ilişki bulunmuştur.

Özdemir (2010)'in belirttiği üzere "Ülkemizde ... bu işle [bankacılık ve finans sektöründe istatistiksel modelleme] ilgilenen analistlerin sayısı artsa da bu konudaki literatür henüz çok sınırlı bir durumdadır." İlgili çalışmada lojistik regresyon analizi kredi skorlama alanında kullanmıştır. Müşteriler, geciken taksit adetlerine göre iyi, kötü ve belirsiz olarak 3 kategoriye ayrılmıştır. Çalışma kapsamında farklı, iyi kötü tanımları yapılarak hangisinin daha iyi bir tanım olduğu tespit edilmeye de çalışılmıştır. Örneklem, 8.603 tane iyi, 2.595 tane kötü ve 2.799 tane belirsiz olarak işaretlenmiş gözlemden oluşmaktadır. Veri seti olarak bir bankanın bireysel portföyünde yer alan müşterilerin başvuru verileri kullanılmış ve bir skor kart modeli geliştirilmiştir.

Anavatan (2011), İş Geliştirme Merkezlerine kayıtlı olan genç firmaların hayatta kalma süresini etkileyen faktörleri Cox regresyon kullanarak analiz etmiştir. Araştırmada kullanılan değişkenler, girişimcinin özelliklerini, firmanın niteliklerini, İş Geliştirme Merkezinde sunulan hizmetleri, ekonomik sektör durumunu ve ekonomik konjonktürü yansıtmaktadır. Bir kuluçka programına kayıt olan firmaların bu kuluçka süresince kapanıp kapanmadığı ve dolayısıyla bir firmanın kurulduktan ne kadar sonra kapandığı modellenmiştir. Bu çalışma ile firmaların hayatta kalmalarına veya kapanmalarına etki eden zamandan bağımsız ve zamana bağımlı değişkenler tespit edilmiştir.

Topcu (2013) imalat sektöründe faaliyet gösteren firmalar bağlamında finansal başarısızlığı açıklayan modeller geliştirmiştir. Finansal başarısızlığın kriteri olarak şirketin üst üste 3 yıl zarar etmesi alınmıştır. İlgili tezde İstanbul Menkul Kıymetler Borsası'na kayıtlı imalat şirketlerinin finansal oranları bağımsız değişken olarak kullanılmış ve sağkalım analizi (zamana bağlı değişkenler içeren Cox regresyon) yapılmıştır. Karşılaştırma maksadıyla aynı veriye lojistik regresyon analizi ve Cox oransal tehlike (hazard) analizi de uygulanmıştır. Bu modellerin çıktıları karşılaştırıldığında sağkalım analizinin daha kesin sonuçlar ürettiği sonucuna varılmıştır.

Kocabaş (2014), lojistik regresyon analizi ile banka müşterilerinin gelecekte temerrüde düşüp düşmeyeceklerini tahmin eden bir model kurmuştur. Çalışmadaki örneklem büyüklüğü 1.000'dir. Bağımlı değişken, müşterilerin güncel ve geçmiş ödeme alışkanlıklarına göre oluşturulmuş ve iyi ve kötü olarak iki kategoriden oluşmaktadır. 7 tane bağımsız değişken kullanılarak müşterilerin ödeme durumlarını etkileyen faktörler belirlenmiştir. Analizde kullanılan bağımsız değişkenler, hesap dengesi, kredi amacı, tasarruf, yaş, medeni durum, cinsiyet ve daire tipidir. Çalışmada, lojistik regresyon analizi ile bir model kurulmuş ve iyi olarak kategorize edilen müşterilerin % 88'i, kötü olarak kategorize edilen müşterilerin % 41,7'si doğru şekilde tahmin edilmiştir.

Cox regresyon analizi ile lojistik regresyon analizini uygulamalı şekilde karşılaştıran makaleler daha çok bioistatistik alanına aittir. Bu tezin kapsamı doğrultusunda finansla ilgili çalışmalara yer verileceği için bioistatistik alanına ait makalelere değinilmeyecektir. Finans ile ilgili belli başlı makaleleri şöyle özetleyebiliriz:

Laitinen ve Kankaanpaa (1999), Finlandiya'daki bir finansal kuruluşa ait veriyi kullanarak aralarında lojistik regresyon analizi ve sağkalım analizinin de olduğu 6 alternatif metot kullanarak finansal başarısızlığı modellemişlerdir. Finansal başarısızlıktan bir yıl, iki yıl ve üç yıl önceki veriler kullanılarak modelleme yapılmıştır. Araştırmanın temel amacı, bu metotlarla elde edilen sonuçların birbirlerinden farklılaşıp farklılaşmadığını analiz etmektir. Analiz sonucunda finansal başarısızlıktan bir yıl önceki verilerle kurulan modelde, lojistik regresyon analizi, sağkalım analizinden daha iyi, finansal başarısızlıktan iki ve üç yıl önceki verilerle kurulan modellerde ise sağkalım analizi, lojistik regresyon analizinden daha iyi olarak elde edilmiştir. Farklı metotlarla tahminlerin doğruluğu açısından farklı sonuçlar elde edilmiş olsa da bu farklar, istatistiksel açıdan önemli değildir. Bu nedenle bu altı yöntemden birisinin diğerlerinden üstün olduğu sonucuna varılamamaktadır.

Stepanova ve Thomas (2001) bir İngiliz finansal kuruma ait veri setine lojistik regresyon analizi ve Cox regresyon analizi uygulamışlar ve iki modelin sonuçlarını ROC eğrisi analizi ile karşılaştırmışlardır. Elde ettikleri sonuca göre, Cox regresyon analizi özellikle kredinin ikinci yılından sonra daha kesin tahminler üretmektedir. Yazarlar, Cox regresyon analizinin, lojistik regresyon analizine bir alternatif olduğunu belirtmişlerdir.

Andreeva (2006), 3 Avrupa ülkesine (Belçika, Hollanda ve Almanya) ait bireysel kredi kartları verilerini içeren bir veri setine çeşitli istatistikî teknikler uygulayarak tahmin modelleri geliştirmiş ve ülke bazında veri setlerini kullanarak model performanslarının ülkeden ülkeye değişip değişmeyeceğini analiz etmiştir. Çalışmada sağkalım analizi ile lojistik regresyon analizinin tahmin gücü açısından yakın sonuçlar verdiği sonucuna varılmıştır. Yazara göre, tahmin güçleri yakın olduğu için sağkalım analizi alternatif diğer faydaları göz önünde bulundurulduğunda daha uygun bir teknik olarak belirmektedir.

Bellotti ve Crook (2009), bir İngiliz bankasının bireysel kredi kartı verilerini içeren bir veri setinde, zamana bağlı değişkenler içeren Cox regresyon analizi ile lojistik regresyon analizi uygulamışlardır. Cox regresyon analizinde konut fiyat endeksi, faiz oranları ve tüketici güven endeksi gibi makroekonomik değişkenler modellemeye dâhil edilmiştir. Bu değişkenlerin eklenmesiyle Cox regresyon analiziyle geliştirilen modelin tahmin gücünde artış olmuştur. Diğer taraftan modelin, makroekonomik değişkenler içermeyen lojistik regresyon modelinin tahmin gücüyle karşılaştırıldığında, makroekonomik değişkenler içeren Cox regresyon analizinin daha kesin tahmin yapmaya imkân sağladığı sonucuna ulaşılmıştır. Yazarlar, Cox regresyon analizinin kredi skorlama açısından lojistik regresyon analizi ile rekabet edebilecek bir rakip olduğunu belirtmişlerdir.

Dominiak ve Mazurkiewicz (2011), Polonya'da yiyecek ve içecek sektöründe faaliyet gösteren firmaların iflas etme riskini Cox regresyon analizi kullanarak modellemişlerdir. Cox regresyon analizi ile kurdukları model sayesinde firmaların iflas durumlarını oldukça kesin bir şekilde tahmin edebilmişlerdir.

Treewichayapong vd. (2011) Tayland'taki gayrimenkul firmaları evreninde iflas tahmin modelleri geliřtirmişlerdir. İlgili veriye lojistik regresyon ve Cox regresyon (Cox Oransal Tehlike modeli) analizi uygulamışlardır. Lojistik regresyon analizi ile kurulan modelin genel tahmin doğruluęu % 95 iken, saękalım analizi ile kurulan modelin genel tahmin doğruluęu % 84,3 olarak hesaplanmıştır. Uygulamanın ardından bu iki modeli tahmin doğruluęu açısından karşılařtırmışlar ve lojistik regresyon analizi ile daha doğru tahminleme yapıldığı sonucuna varmışlardır.

Im vd. (2012) bir Amerikan finansal kuruluşunun veri setini kullanarak tüketici kredileri portföyünde saękalım analizi ile modelleme yapmışlardır. Çalışmada saękalım analizinin lojistik regresyon analizine bir alternatif olduęu belirtildikten sonra saękalım analizinin sunduęu 3 avantajdan bahsedilmiştir. Bunlardan birincisi, saękalım analizinde, temerrüt zamanına ait dağılım modellenir. Böylece çok farklı zaman dilimleri için (örneğin 12 aylık veya 24 aylık) temerrüt tahminleri yapılabilir. İkincisi, en güncel verinin kullanılması mümkündür. Şöyle ki lojistik regresyon analizinde öncelikle bir performans dönemi belirlenir ve tüm gözlemlerin performans dönemi süresince gözlemlenmesi gerekir. Örneğin performans dönemi olarak bir yıl alındığında Ocak 2018'de yapılacak bir çalışmada en güncel veri Ocak 2017'ye ait olabilir. Saękalım analizinde ise Ocak 2017 ile Ocak 2018 arası veri analize dâhil edilebilir. Üçüncüsü, saękalım analizi daha kesin bilgi vermektedir. Diğer bir ifade ile müşterinin temerrüt edip etmeyeceęi bilgisinin yanı sıra ne zaman temerrüt edeceęi bilgisini de içerir.

Louzada vd. (2014) saękalım analizinin kredi riski modellemesinde kullanılmaya başlanmasını ve yaygınlaşmasını bankaların sermaye yeterlilięi hesaplamasına ilişkin Basel standartları bağlamında ele almıştır. Sermaye yeterlilięi hesaplamasında içsel derecelendirmeye dayalı yöntemle izin verilmesiyle bankaların kredi riski bileşenlerinden birisi olan temerrüt olasılıęını modelleme ihtiyacı daha da belirginleşmiştir. Bu bağlamda saękalım analizi lojistik regresyon analizine göre daha fazla bilgi verdięi için kullanılmaya başlanmıştır. Saękalım

analizi, müşterinin temerrüt edip etmeyeceğinin yanı sıra ne zaman temerrüt edeceği bilgisini de sağladığı için kredi riski yönetimi açısından da daha kullanışlıdır. Bu özelliği de onun yaygınlaşmasına katkıda bulunmuştur. İlgili çalışmada, sağkalım analizinin, oransal tehlike (hazard) varsayımı sağlanmadığı durumlarda, üretilen sonuçların çok yanlış olabileceğine vurgu yapılmıştır.

Ptak-Chmielewska (2016), bir Polonya bankasının verisini kullanarak içsel derecelendirme modelleri bağlamında aralarında lojistik regresyon analizi ve Cox regresyon analizinin de olduğu 5 farklı istatistiksel model geliştirmiştir. Çalışmada her bir modelin avantaj ve dezavantajlarını anlattıktan sonra, Birinci Tip Hata (temerrüt edenlerin yanlış bir şekilde canlı olarak sınıflandırılması), İkinci Tip Hata (canlı müşterilerin yanlış bir şekilde temerrüt olarak sınıflandırılması) ve GINI Katsayısı açısından modelleri karşılaştırmıştır. GINI Katsayısı açısından lojistik regresyon analizi ve Cox regresyon analizinin sonuçları birbirine çok yakındır. Cox regresyon analizi ile geliştirilen model Birinci Tip Hata açısından daha güvenilir sonuç verirken İkinci Tip Hata açısından lojistik regresyon analizi daha doğru sonuç vermiştir. Yazar, içsel derecelendirme modelleri bağlamında lojistik regresyon analizinin yaygın olarak kullanıldığına vurgu yaptıktan sonra Cox regresyon analizinin kullanımının da giderek arttığını belirtmiştir.

Marimo ve Chimedza (2017), önde gelen bir Güney Afrika finansal kuruluşuna ait veriyi kullanarak hem lojistik regresyon modeli hem de Cox regresyonu modeli geliştirmişlerdir. Veri 2009 ile 2014 yılları arası döneme aittir. Modellerin sağlamlığını arttırmak için değişkenlere WoE dönüşümü uygulanmıştır. Veri, öğrenme ve doğrulama verileri olarak ikiye ayrılmıştır. Kurulan modeller GINI katsayısı ve ROC Eğrisi altındaki alan baz alınarak karşılaştırılmıştır. Cox regresyon modeli her iki kriter açısından da daha iyi performans göstermiştir.

Genel olarak ifade etmek gerekirse, bu çalışmalarda genellikle spesifik bir alana (örneğin belli bir ekonomik sektörde faaliyet gösteren firmalar) odaklanılmış ve sonuçlar, ancak o alanda genelleştirilmiştir. Bu tezde ise farklı bir spesifik

alana odaklanılmıştır. Bundan dolayı bu arařtırmaların sonuçları ile banka müşterileri üzerine yapacağımız çalışmanın sonuçlarının doğrudan karşılaştırılması doğru olmayacaktır. Çünkü birden fazla banka ile çalışan müşterilerin geri ödeme davranışları ve çalışma şekilleri bankadan bankaya değişebilmektedir. Aynı müşterinin farklı bankalarda farklı şekilde bilgileri bulunabilmektedir. Bu yüzden bir bankanın verisi ile geliştirilen bir model başka bir bankada aynen kullanılamaz. Bu çalışmalar örneklem büyüklüğünün belirlenmesinde ve tahminlerde kullanılan değişkenlerin tespitinde oldukça faydalıdır.



3. TFRS 9 VE TEMERRÜT OLASILIĞI MODELLEMESİNDEKİ KAVRAMLAR

Çalışmanın bu bölümünde tezde kullanılan kavramlar, tez konusunun daha iyi anlaşılabilmesi açısından anlatılmaya çalışılmıştır.

3.1 Beklenen Kredi Zararı

Beklenen kredi zararı, ilgili mevzuatta detaylıca anlatılmış ve çalışmamızda bu tanımlara yer verilmiştir. TFRS 9'a göre beklenen kredi zararı (BKZ), "finansal aracın beklenen ömrü boyunca kredi zararlarının olasılıklarına göre ağırlıklandırılmış bir tahminidir." Kredinin aşamasına göre beklenen kredi zararı hesaplaması değişecektir. Kredi aşamasına göre beklenen kredi zararı, 12 aylık ve ömür boyu olmak üzere iki şekilde hesaplanacaktır.

TFRS 9'a göre; 12 aylık beklenen kredi zararı, "Ömür boyu beklenen kredi zararlarının, raporlama tarihinden sonraki 12 ay içinde finansal araca ilişkin gerçekleşmesi mümkün temerrüt hallerinden kaynaklanan beklenen kredi zararlarını temsil eden kısmıdır." TFRS 9'a göre; ömür boyu beklenen kredi zararı ise, "Finansal aracın beklenen ömrü boyunca gerçekleşmesi mümkün tüm temerrüt durumlarından kaynaklanan beklenen kredi zararlarıdır."

TFRS 9'a göre; "Beklenen kredi zararlarının tahmin edilmesindeki amaç, ... beklenen kredi zararlarının tahmini, her zaman kredi zararının meydana gelme olasılığı ile kredi zararı meydana gelmeme olasılığını yansıtır" ve "tahmini beklenen kredi zararlarının, mümkün sonuç aralığı değerlendirilerek belirlenen tarafsız ve olasılıklara göre ağırlıklandırılmış tutarı yansıtmasını zorunlu kılar."

Beklenen kredi zararı hesaplaması, 2018 öncesi mevzuatta yer alan özel karşılık ve genel karşılık hesaplamaları yerine kullanılmaktadır. Her bir müşteri ve

teminat için aynı sabit oranların kullanıldığı genel ve özel karşılık ayırımı ortadan kalkmıştır.

Aytürk (2016) aşamalar arası geçişi ve beklenen kredi zararı hesaplamasını genel hatlarıyla ele almıştır. Onun da belirttiği üzere, “işletmelerin içsel kredi derecelendirme sistemlerindeki verileri kullanarak kredi riski modellemesi yaparak her bir kredi için beklenen zarar oranını hesaplamaları gerekmektedir.” Bankaların, bu yeni yaklaşım gereği, tüm müşterileri için tekil veya havuz bazında aşağıda belirtilen bileşenlere dair modeller kurması elzemdir. Bankacılık uygulamaları açısından Beklenen Kredi Zararı yaklaşımı 4 ana bileşenden oluşmakta ve genel hatlarıyla şöyle ifade edilmektedir:

Beklenen Kredi Zararı = Temerrüt Olasılığı * Temerrüt Halinde Kayıp * Temerrüt Anındaki Risk * İskonto Oranı.

Her bir bileşen ayrı birer tezin konusu olacak kadar kapsamlı konulardır. Bu tezin kapsamı dâhilinde sadece temerrüt olasılığı modellemesi üzerinde durulacaktır.

3.2 Finansal Araçlar Raporlama Standardı

TRFS 9; Sınıflama ve Ölçüm, Değer Düşüklüğü ve Riskten Korunma Muhasebesi olmak üzere 3 ana bölümden oluşur ve “finansal tablo kullanıcılarına işletmenin gelecekteki nakit akışlarının tutarını, zamanlamasını ve belirsizliğini değerlendirmeleri için ihtiyaca uygun ve faydalı bilgiyi sunacak şekilde finansal varlıklara ve finansal yükümlülüklerle ilişkin finansal raporlama ilkelerini belirle[r].” (Kamu Gözetim Kurumu, 2018). Bu standart Türkiye’ye özgü olmayıp ilgili uluslararası mevzuatın yani International Financial Reporting Standards 9 (IFRS 9)’un Türkiye mevzuatına kazandırılmış halidir (IFRS 9’un tarihi arka planı, etkileri ve yaygınlaşması hususlarında bkz., De George vd, 2016). Kamu Gözetimi Muhasebe ve Denetim Standartları Kurumu (KGGK)’nun TRFS 9’a dair tebliği 19 Ocak 2017’de Resmi Gazete’de yayınlanmıştır. Yürürlük tarihi olan 1 Ocak 2018’ten beri uygulanmaktadır.

Uluslararası boyut üzerinde kısaca durmak gerekirse, kâr amacı gütmeyen ve kamu yararı için çalışan bir örgüt olan UFRS Vakfı (IFRS Foundation), anlaşılabilir, kaliteli, uygulanabilir ve dünya çapında genel kabul görmüş ortak muhasebe standartları geliştirmek ve o standartların benimsenmesini teşvik etmek ve kolaylaştırmak için kurulmuştur. Yayınladığı standartlar IFRS Standartları olarak adlandırılır. Bu standartlar, vakfin bir kurulu olan Uluslararası Muhasebe Standartları Kurulu tarafından geliştirilip yayınlanır. IFRS 9 standardı da bu kurul tarafından yayınlanmıştır. Başta G-20 ülkeleri olmak üzere, Dünya Bankası gibi birçok uluslararası örgüt ve ülke, ortak muhasebe standartlarının geliştirilmesini desteklemektedir. (UFRS Vakfı, 2018).

Bu tezde bir bütün olarak TFRS 9 ele alınmayacak, ancak TO modellemesi ile doğrudan irtibatlı olan Değer Düşüklüğü bölümü üzerinde durulacaktır. Standardın bu bölümünde detaylandırıldığı üzere TFRS 9, gerçekleşen zarar yaklaşımı yerine beklenen zarar yaklaşımını esas almaktadır.

TFRS 9 öncesi uygulamada olan gerçekleşen zarar yaklaşımında, canlı kredi olarak kabul edilen ve 1. Grup veya 2. Grup olarak sınıflandırılan kredilere, mevzuat tarafından belirlenen standart oranlarda genel karşılık ayrılıyordu. 3. Grup, 4. Grup ve 5. Grup olarak sınıflandırılan kredilerde, diğer bir ifadeyle donuk alacaklarda temerrüt gerçekleştikten sonra, mevzuat tarafından belirlenen standart oranlarda özel karşılık ayrılıyordu. Karşılık hesaplaması mevzuat tarafından belirlenmiş oranlar üzerinden ve donuk alacaklar için temerrüt olayı gerçekleştikten sonra yapıldığı için, çok kolay bir şekilde yapılabilirdi. Beklenen zarar yaklaşımında ise bankaların temerrüt olayı gerçekleşmeden önce her bir müşterisi için temerrüt olasılığını hesaplaması ve bu hesaplanan olasılıkların karşılık hesaplamasında girdi olarak kullanılması gerekmektedir. Eski uygulamada 1. Grupta yer alan tüm müşteriler için aynı standart oran kullanılırken yeni uygulamada her bir müşteri için ona özgü hesaplanan temerrüt olasılığı kullanılmaktadır. Bu yüzden müşterilere ayrılan karşılıklar oransal olarak ciddi olarak farklılaşabilmektedir. Yeni uygulamada her bir kredi için

raporlama dönemlerinde 12 aylık beklenen kredi zararının veya ömür boyu beklenen kredi zararının hesaplanması ve bilançoya yansıtılması gerekmektedir. 12 aylık beklenen kredi zararı veya ömür boyu beklenen kredi zararının hangisinin kullanılacağı, müşterinin aşamasına göre belirlenmektedir.

3.3 Kredi Riski

Bir müşteriye kredi kullanıldığında, müşterinin borcunu zamanında ve eksiksiz olarak geri ödeyeceği kesin değildir, yani bir belirsizlik söz konusudur. Bu belirsizliği kredi riski olarak ifade edebiliriz. Kredi riski, en genel anlatımla, kredi kullanan bir müşterinin bankaya olan geri ödeme yükümlülüklerini yerine getirememesi ihtimalinden kaynaklanan risktir. Kredi riski ölçümünde kullanılan modeller, “göreceli olarak niteliksel olan modellerden yoğun biçimde nicel olabilen modellere kadar çok geniş bir çerçevede” geliştirilebilmektedir (Kavcıoğlu makalesinde bu modelleri genel hatlarıyla tanıtmıştır. Bkz., Kavcıoğlu, 2011).

TFRS 9’a göre bankaların kredi riskinde önemli artış olan kredileri tespit etmeleri ve ikinci aşamaya alarak ömür boyu beklenen kredi zararı ayırmaları gerekmektedir. Kredi riskinde önemli artış, içsel derecelendirme notunda bozulma veya TO değerlerindeki kötü yönde değişim üzerinden ölçülebilir. Bunun için kredinin verildiği tarihte müşteri için hesaplanan temerrüt olasılığı ile ilgili raporlama dönemi itibarıyla hesaplanan temerrüt olasılığı karşılaştırılabilir. Belirlenen eşik değerden -örneğin % 10 kötüleşme gibi- daha fazla kötüleşme varsa kredi riskinde önemli artış olduğu kabul edilir.

Kredi riski yönetiminin önemli bir bileşeni, her bir müşteri için temerrüt olasılığı üreten içsel derecelendirme modellerinin geliştirilmesi ve çıktılarının banka genelinde ana bankacılık faaliyetlerinde etkin bir şekilde kullanılmasıdır. Böylece bankanın maruz kaldığı kredi riski ve kredi kalitesi ölçülebilir ve izlenebilir. Eski mevzuatta karşılık ayırmada standart oranlar uygulandığı için risk yönetimi uygulamaları karşılık muhasebesine doğrudan etki etmiyordu. Saltoğlu (2016:

55)'nün belirttiđi üzere; TFRS 9, eski standardın “sorunlu taraflarını ortadan kaldırarak bankaların kredi riski yönetimi uygulamaları ile kredi zararlarının muhasebesini uyumlu hale getirmeyi hedeflemektedir.”

3.4 Temerrüt Olasılığı

Temerrüt, bir müşterinin kullandığı kredisini zamanında geri ödeyememesi ve 90 günü aşan gecikmesidir. Genelde temerrüdün esas belirleyicisi 90 günü aşan gecikmedir. Bir temerrüt olasılığı modeliyle bir borçlunun belirli bir süre zarfında temerrüt edip etmeyeceđi tahmin edilir. Bu süre çođunlukla 12 aydır. Temerrüt olasılığı modellemesinde başvuru modeli kurulması ve davranışsal model geliştirilmesi gibi iki temel uygulama mevcuttur. Bu iki yaklaşım birbirinin alternatifi değildir, farklı kullanım alanları vardır.

Başvuru modelleri, ilk defa bankaya başvuran ve bankanın tam anlamıyla tanımadığı müşterilere yöneliktir. Bu yüzden müşterinin başvurduğu bankadaki kredi geri ödeme alışkanlıkları ile ilgili iç kaynaklara dayanan bir tahsilat verisine yer verilememektedir. Ancak bu modellerde KKB, Risk Merkezi (eski memzuç) gibi dış kaynaklı ödeme davranış bilgilerine dayanan verilere yer verilmektedir. Davranışsal modeller ise, başvuru modelinde yer alan verilere ilaveten, müşterinin çalıştığı bankadaki ödeme davranışına ait verileri de içermektedir. Çođu firma aynı anda birden fazla banka ile çalışmakta ve ödeme davranışları bankadan bankaya göre deđişebilmektedir. Davranışsal modeller, iç kaynaklara dayanan ödeme verisini/performansını içerdđi için müşteri hakkında daha net bilgi sağlayabilmektedirler.

Bankalar kredi için başvuran müşterilerinin bir kısmının kredi talebini ret ederler. Bu kitleye kredi tahsisi yapılmadığı için sonraki 12 ayda temerrüt edip etmeyecekleri hususunda bir bilgi mevcut değildir. Yine de başvuru modellemesinde bu ret edilen kitlenin adedinin ve oranının modellemede dikkate alınması gerekmektedir. Başvuru modelleri bu açıdan da davranış modellerinde ayrılır.

Bu alıřmada bankanın tanımadığı ve ilk defa limit tahsisi yapacağı müşteriler hedef kitleyi oluřturmamaktadır. Modellemeye konu müşteri kitlesi, banka ile kredili alıřıyor olan kurumsal müşterilerdir. Dolayısıyla davranıřsal bir model geliřtirilecektir.



4. TFRS 9 VE TEMERRÜT OLASILIĞI MODELLEMESİNDE KULLANILAN İSTATİSTİKSEL YÖNTEMLER

Bir bağımlı değişkenin alacağı değeri tahmin etmek için çok çeşitli istatistiksel teknikler mevcuttur. Doğrusal regresyon, lojistik regresyon analizi, diskriminant analizi, yapay sinir ağları ve sağkalım analizi bunlardan birkaçıdır. Bununla birlikte, Lojistik regresyon analizi, bankacılık alanında en çok kullanılan yöntemlerden birisidir. Cox regresyon analizi ise bankacılık alanında nispeten yeni kullanılmaya başlanmıştır. Bu çalışmada kullanılacak olan istatistiksel teknikler, lojistik regresyon analizi ve Cox regresyon analizidir.

4.1 Lojistik Regresyon Analizi

Lojistik regresyon analizinde bağımlı (yordanan) ve bağımsız (yordayıcı) değişkenler arasındaki ilişki modellenir ve kategorik bağımlı değişkenin aldığı değerlere ilişkin bir üyelik tahmini yapılır (Lojistik regresyon analizinin detaylı anlatımı için bkz., Harrell, 2001; Kleinbaum ve Klein, 2002; Agresti, 2013). Lojistik regresyon analizi, kendi içinde bağımlı değişkenin sahip olduğu kategori sayısı ve kategoriler arasındaki ilişkiye göre, ikili (binary), multinomial (multinomial) ve sıralı (ordinal) olmak üzere üçe ayrılır. Lojistik regresyon analizi denildiğinde aksi belirtilmediği sürece, ikili lojistik regresyon analizi anlaşılmaktadır. Bu tezde de aynı yaklaşım benimsenmiştir.

İkili lojistik regresyon analizinde, bağımlı değişkenin iki kategorisi vardır. Örneğin, temerrüt etmiş – temerrüt etmemiş, başarılı – başarısız gibi iki değer alır. Multinomial Lojistik Regresyon (MLR) analizi ise, ikili lojistik regresyonun bağımlı değişkenin ikiden fazla kategorisi olduğu durumlar için geliştirilmiş halidir (Hosmer vd, 2013: 269). Bağımlı değişkenin üç veya daha fazla kategoriden oluştuğu ve kategoriler arasında bir hiyerarşi olmadığı durumlarda MLR uygulanır. MLR'nin en önemli özelliği bağımlı değişkenin bu çok kategorili

olma durumudur. MLR'de bağımsız değişkenin bir kategorisi referans kategori olarak seçilir ve diğer kategoriler bu referans kategoriye göre yorumlanıp analiz edilir (MLR'nin nasıl yapıldığı hususunda bkz., Petrucci, 2009; Kleinbaum ve Klein, 2002). Sıralı lojistik regresyon analizinde ise, bağımsız değişken gene en az 3 kategoriden oluşur, ancak burada kategoriler arasında üstünlük-altlık yani bir hiyerarşi vardır.

Lojistik regresyon analizinin matematiksel denklemleri (detaylı bir anlatım için bkz., Kleinbaum ve Klein: 2002) aşağıda kısaca izah edilmiştir. Lojistik modelin temelinde z hangi değeri alırsa alsın 0 ile 1 arasında sonuç üreten lojistik fonksiyonu vardır. Bu fonksiyon;

$$f(z) = \frac{1}{1 + e^{-z}} \quad (4.1)$$

olarak ifade edilir.

Lojistik regresyon analizinde bir bağımlı (Y ile ifade edilmiş olsun ve olayın gerçekleşmesini 1, olayın gerçekleşmemesini ise 0 temsil ediyor olsun) ve bir bağımsız değişken (X ile ifade edilmiş olsun) olduğu durumda basit lojistik regresyon modeli kurulur ve model;

$$\pi(x) = P(Y = 1|X = x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1)}} \quad (4.2)$$

olarak ifade edilir.

İki ve daha fazla bağımsız değişken olduğu durumlarda çoklu lojistik regresyon modeli kurulur ve çoklu lojistik regresyon modeli;

$$\pi(x) = P(Y = 1|X = x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad (4.3)$$

olarak ifade edilir.

4.1.1 Lojistik regresyon analizinin varsayımları

Her istatistiksel model bazı varsayımlara dayanmaktadır. Model kurmaya başlamadan önce bu varsayımların karşılanıp karşılanmadığı sınıanmalıdır. Böylece kullanılmak istenen regresyon analizinin doğru analiz türü olup olmadığına kesin şekilde karar verilebilir. Varsayımlar karşılanmadığı durumlarda ya mevcut veri setine uygun başka bir analiz yöntemi uygulanmalı ya da veri setinde bazı dönüşümler yapılarak analize uygun hale getirilmelidir. Varsayımlar karşılanmadığı durumlarda elde edilen sonuçlar, geçerli sonuçlar olmayacaktır. Veri seti, ilgili varsayımları karşılamadığı halde karşıladığı varsayımları analizler yapıldığında ulaşılan sonuçların uygulama değeri de olmayacaktır. Örneğin bu çalışmada, varsayımlar karşılanmadan bir model kurulsa, uygulamada hiçbir banka karar alırken bu modeli kullanmak istemeyecektir.

Lojistik regresyon analizinin varsayımları (Bu hususta detay için bkz., Tabachnick ve Fidel, 2015) ve bu varsayımlara dair yapılan incelemeler aşağıda görüldüğü gibidir.

Birinci varsayım, binom dağılım şartının sağlanmasıdır. İki sonuçlu olayların gerçekleşme olasılıklarına ilişkin dağılım binom dağılımı olarak adlandırılmaktadır. Binom dağılımının sağlanması için gerekli olan şartları şöyle özetleyebiliriz: Birincisi, lojistik regresyon analizinde olduğu gibi incelenen olay iki seçenekli olmalıdır. Verilen kredi ya temerrüt eder ya da etmez. İkincisi, her bir olay birbirlerinden bağımsız olmalıdır. Üçüncüsü, iki seçenekten birisinin gerçekleşme olasılığı her bir denemede değişmemelidir (Turanlı vd., 2017: 72-74).

Peng vd. (2002: 11)'nin belirttiği üzere, örneklem rassal olduğu ve gözlemler eşdeğer ve birbirlerinden bağımsız olduğunda, literatürde bu varsayımın sağlandığı varsayılmaktadır. Bu tez çalışmasında binom dağılımının temel teorik

varsayımları sağlandığı için ilave bir analiz yapılmamış ve bu varsayımın sağlandığı kabul edilmiştir.

İkinci varsayım, gözlemler birbirlerinden bağımsız olmalıdır. Gözlemler, tekrarlanan ölçümlerden oluşmamalıdır. Lojistik regresyon analizinin de dahil olduğu çoğu istatistiksel test, gözlemlerin birbirinden bağımsız olması temel varsayımı üzerine bina edilmiştir. McDonald (2014: 131-2)'ın belirttiği üzere, gözlemler arasında bağımlılık mevcutsa, Birinci Tip hata (false positive) oranı, diğer bir ifadeyle temerrüt etmiş gözlemlerin temerrüt etmemiş olarak tahmin edilmesi yüksek çıkabilir.

Gözlemlerin birbirinden bağımsız olmasını biraz daha açmak faydalı olacaktır. Bir gözlem, başka bir gözlem hususunda doğrudan bir bilgi sunuyor olmamalıdır. Örneğin 0-1 gibi ikili bir yapı olduğunda, bir gözlemin 1 olarak gerçekleşmesi başka bir gözlemin 1 olarak gerçekleşme ihtimalini azaltmamalı veya arttırmamalıdır. Bazı gözlemler arasında bağımlılık söz konusu ise rassallık ilkesi ihlal edilmiş olmaktadır.

McDonald (2014: 131-2)'ın belirttiği üzere, gözlemlerin birbirlerinden bağımsız olmayışının sebebi çoğunlukla gözlemlerin mekân ve zaman açısından birbirine çok yakın olmasıdır. Bu doğrultuda aylık veya üç aylık aralıklar yerine altışar aylık aralıklarla oluşturulmuş veri ile modelleme çalışması yapılmıştır. Yine de altışar aylık periyotlarla veri oluşturulduğu için bir müşterinin bütün dönemlerde olması veya aynı temerrüt olayı ile iki defa temerrüt olarak veriye girmesi mümkündür. Böyle durumlarda ilgili gözlemler çıkarıldığında müşterinin sonraki performansı hakkında bilgi kaybının olmadığı gözlemler elenmiştir ve aynı bilginin iki defa veride yer almasının önüne geçilmiş ve böylece gözlemlerin birbirinden bağımsız olmasının sağlanmasına destek olunmuştur.

Üçüncü varsayım, örneklem yeteri kadar büyük olmalıdır. Lojistik regresyonda parametre tahmininde en çok olabilirlik yöntemi kullanıldığı ve bu yöntem büyük örneklem gerektirdiği için mümkün oldukça büyük örneklem ile

çalıřılması önerilmektedir. Örneklem büyüklüğü hususunda Çokluk (2010: 1361)'un önerisi şöyledir: Bağımlı deęişken ikiden fazla kategoriden oluşuyorsa her bir bağımsız deęişken için minimum 50 kişilik bir grup büyüklüğüne gerek vardır. Alpar (2017: 601), deęişken sayısını baz alarak en az 10 katı gözlemle analiz yapılması gerektiğini belirtmektedir.

Tahmin gücü yüksek ve güvenilir bir modelin kurulabilmesi için yeteri kadar temerrüt etmiş müşteri olması gerekmektedir. Örneklem büyüklüğü hususunda karşılaşılan en büyük sorun yeteri kadar temerrüt eden müşteri olmaması durumudur. Bu konuda tam bir standart olmasa da piyasada iyi uygulama kriteri olarak baz alınan bazı eşik deęerler mevcuttur. Örneğin Sıddıqı (2017: 75) en az 2000 tane temerrüt etmiş müşterinin yer alması gerektiğini ifade eder.

Modelin geliştirileceęi portföyde piyasa uygulamaları baz alındığında yeteri kadar temerrüt etmiş müşteri bulunmaktadır.

Dördüncü varsayım, çoklu doğrusallık (multicollinearity) olmaması gerekir. Lojistik regresyon bağımsız deęişkenler arasındaki yüksek korelasyona oldukça duyarlıdır. Aralarında çoklu doğrusallık olan deęişkenlerin olması durumunda deęişkenlerin azaltılması yani sorunu çözmeye yetecek kadar deęişkenin çıkarılması gerekir.

Çoklu doğrusallık sorunu, bağımsız deęişkenlerin birbirleriyle güçlü şekilde ilişkili olmasıdır. Çoklu regresyon denkleminin temelinde, bağımsız deęişkenlerin kuvvetli bir şekilde ilişkili olmadığı varsayımı bulunmaktadır. Bu varsayım sağlanamadığında denklemin yorumlanması problemlile hale gelmektedir. Alpar'ın (2017: 505, 506) belirttięi üzere, çoklu doğrusallık sorunu varsa "regresyon modeli yardımıyla yapılacak çıkarsamalar yanlış yönlendirmelere ve hatalara neden olur. Örneğin veride çok [bir] küçük deęişiklik yapıldığında katsayılarda büyük deęişiklikler olur. ... Regresyon sonuçlarında belirsizlikler söz konusu olacaktır. ... elde edilen regresyon

katsayıları, ... modelden bir değişkenin çıkartılması ve eklenmesine karşı çok duyarlı hale gelir.”

Çoklu doğrusallık sorununun çözümü için faktör analizi yöntemiyle modeldeki değişkenler birleştirilip elde edilen faktörlerle model kurulabilir veya bu soruna neden olan değişkenler modelden çıkartılabilir. Doğrusal regresyonda bu sorunun tespiti için çeşitli hesaplamalar yapılabilir (Çokluk vd, 2014: 35-38).

Öncelikle, değişkenler arasındaki basit korelasyon incelenir. Değişkenler arasında güçlü korelasyon olması çoklu doğrusallık sorununa işaret eder. Daha sonra, VIF (Variance Inflation Factor = Varyans artış faktörü) incelenir. Genel kural olarak VIF'in 10'a eşit veya daha büyük olması durumunda çoklu bağlantı sorunu olduğu anlaşılmış olur. Ardından, tolerans değeri incelenir. Bu değer 0.1'den büyük olması durumunda, değişkenler arasında çoklu doğrusallık sorununun olmadığı söylenebilir. Daha sonra, Condition Index incelenir. Bu değer 10 ile 30 arasında olması orta düzeyde, 30'dan büyük olması durumunda yüksek düzeyde çoklu doğrusallık problemi olduğu anlaşılır.

Lojistik regresyon modelinde bağımsız değişkenler arasında çoklu doğrusallık varsa “katsayıların standart hataları olduğundan çok büyük kestirilir. ... regresyon katsayılarının standart hatası büyük kestirildiğinden Wald istatistiği küçülmekte ve önemli bir değişkenin etkisi önemsiz bulunmaktadır. ... regresyon katsayıları da çoklu bağlantı nedeniyle aşırı büyük kestirilme eğilimindedir.” (Alpar, 2017: 613). Bu yüzden katsayıları ve standart hataları büyük olan değişkenler muhakkak analiz edilmeli ve çoklu doğrusallık olup olmadığı analiz edilmelidir.

Modellemeye konu değişkenler arasında yüksek korelasyon tespit edilmediğinden dolayı, çoklu doğrusallık tespiti için ilave analiz yapılmamış ve çoklu doğrusallık sorunu olmadığı varsayılmıştır.

Beşinci varsayım, lojistik model, uç değerlere son derece duyarlıdır. Bağımsız değişkenlerdeki uç değerlin iyi analiz edilmesi gerekir. Çoğu istatistiksel işlemin temelinde ortalamadan sapmaların karesi yer almaktadır. Eğer bir gözlem ortalamadan çok uzakta ise sapma değeri büyüyecektir.

Öncelikle banka tarafından veri kontrolleri yapılırken uç değerler açısından da kontrol yapılmıştır. Veri girişi yapanların hatalarından kaynaklı hatalar tespit edildiğinde kaynak tablolarında düzeltme yapılmıştır. Modelleme safhasında her bir değişkendeki uç değerler, değişken özelinde analiz edilmiş ve ilgili iş birimleriyle görüşülüp bilgi alınmıştır. Uç değerlerin etkisini azaltmanın bir yolu, sürekli değişkenleri yeniden sınıflandırırken uç değerler için ayrı bir kategori oluşturulmasıdır. Böylece analizde, ilgili gözlem analiz dışında bırakılmadığından bilgi kaybı minimize edilmektedir.

Çalışmamızda, modellemede kullanılan sürekli değişkenler, monotonik yapıyı sağlayacak şekilde yeniden sınıflandırılırken, uç değerler için ayrı birer kategori oluşturularak veya bu uç değerler sonuncu kategoriye dâhil edilerek, bu sorun minimize edilmiştir.

Altıncı varsayım, veri analize başlanmadan önce kayıp değerler açısından kontrol edilmeli ve kayıp değerlere nasıl bir uygulamanın yapılacağı belirlenmelidir.

Kayıp değerler hususunda, hangi örneklem büyüklüğünde ne kadar kayıp değer göz ardı edilebilir olduğuna ilişkin literatürde belirlenmiş bir ölçüt yoktur. Çözümlerden birisi, % 40 gibi belli bir oranın üzerinde kayıp değer içeren değişkenlerin modelden çıkartılmasıdır. Diğer bir çözüm, kayıp değer yerine ortalama değeri yazmaktır. Üçüncü bir yöntem, sürekli değişkenler yeniden sınıflandırılırken kayıp değerler için ayrı bir kategori oluşturulmasıdır. Bunun şöyle bir faydası olabilmektedir. Bazı durumlarda kayıp değerlerin kendisi ayrı bir bilgi sunuyor olabilir. Bu yüzden kayıp değerler her bir değişken özelinde değerlendirilmeli ve mümkün olduğunca modellemeye dahil edilmelidirler.

Bu çalışmada kayıp değer yüzdesi çok yüksek olan değişkenler çıkartılmış, tolere edilebilir düzeyde kayıp değer olduğu durumlarda, kayıp değerlere özgü kategoriler oluşturulmuştur.

Yedinci varsayım, uygun tüm bağımsız değişkenler modele dâhil edilmelidir. Açıklayıcılık gücü olan değişkenlerin modele dâhil edilmemesi modelin yetersizliğine neden olabilir.

Alpar (2003: 192)'ın belirttiği üzere araştırmada kullanılacak bağımsız değişkenler tespit edilirken karşılaşılan önemli bir sorun, belirleme hatası yani "ilgisiz değişkenlerin modele alınması, ilgili değişkenlerin ise model dışında bırakılmasıdır". İlgisiz değişkenlerin elenmesi hususunda Alpar (2017: 611-2) iki seçenek önerir. Birinci seçenek, bağımsız değişkenin bağımlı değişkenle ilişkili olup olmadığının ki-kare, iki ortalama arası farkın önemlilik testi, Mann-Whitney U gibi testler aracılığıyla belirlenmesidir. Bu analizlerle ilişkisiz olduğu tespit edilen bağımsız değişkenler modelleme çalışmasının kapsamından çıkartılır. İkinci seçenek, her bir bağımsız değişken ile birer lojistik regresyon analizi yapılarak, tek değişkenli lojistik regresyon analizi ile elde edilen p değerleri analiz edilir ve p değeri 0,25'ten küçük olan bağımsız değişkenler modellemeye alınır.

Modellemede temel amaç modellemeye konu ilişkiyi en az sayıda değişkenle açıklamaktır. Bu yüzden ilgisiz değişkenlerin öncelikle elenmesi gerekmektedir. Diğer taraftan, ilgili bir değişken modellemeye alınmadığında ise modelin açıklayıcılığı azalacaktır.

Çalışmada, Sıddıqî'nin yaklaşımı ile ön analizler yapılarak ilgili değişkenlerin modellemede yer alması sağlanmıştır.

Sekizinci varsayım, sürekli bağımsız değişkenler ile bağımlı değişkenin logit dönüşümü arasında doğrusal bir ilişki olmalıdır.

Bağımlı değişkenin logit dönüştürmeye tabi tutulmuş hali ile sürekli bağımsız değişkenler arasında doğrusal ilişki olması gerekir. Alpar (2017: 603) doğrusal ilişkinin olup olmadığını test etmek için Box-Tidwell test yaklaşımının sıklıkla kullanıldığından bahseder ve bu yaklaşımı şöyle özetler:

“sayısal bağımsız değişken x ile bu değişkenin doğal logaritması $\ln(x)$ çarpılarak, $x\ln(x)$ etkileşim terimi modele eklenir. Etkileşim teriminin modelde önemli bulunması lojit ile sayısal bağımsız değişken arasındaki ilişkinin doğrusal olmadığını belirtir. ... Varsayım sağlanmadığında, bağımsız değişkende dönüşüm yapılarak ilişki... doğrusallaştırıl[abilir] ya da sayısal değişken... sınıflanarak niteliksel değişken olarak modele alın[abilir.]”

Çalışmadaki tüm sürekli değişkenlere dönüşüm uygulanarak, monotonik hale getirilmiş ve modellemede değişkenlerin dönüştürülmüş halleri kullanılmıştır. Uygulanan dönüşüm, uygulama bölümünde etraflıca anlatılmıştır.

4.1.2 Lojistik regresyon analizinin avantajları ve kısıtları

Lojistik regresyon analizi, diğer regresyon türleriyle karşılaştırıldığında (detaylı bir karşılaştırma için bkz., Karagöz, 2016) varsayımlar açısından nispeten esnek bir yapıya sahiptir. Bağımlı değişkenin kategorik yapıda olması ve bağımsız değişkenler için normal dağılım şartının aranmaması uygulanmasını kolaylaştırmaktadır. Diğer taraftan olasılıklar 0 ile 1 aralığında olduğundan sonuçlar kolaylıkla yorumlanabilmektedir. Bu nedenle, tıptan bankacılığa kadar birçok alanda kullanılabilir. Bu nedenle, tıptan bankacılığa kadar birçok alanda kullanılabilir.

Alpar (2017: 600) lojistik regresyon analizinin tercih edilme nedenlerini “fonksiyonun matematiksel açıdan çok esnek olması, kullanım kolaylığı ve sonuçların klinik açıdan anlaşılabilir bir şekilde yorumlanabilmesi” olarak belirtmektedir.

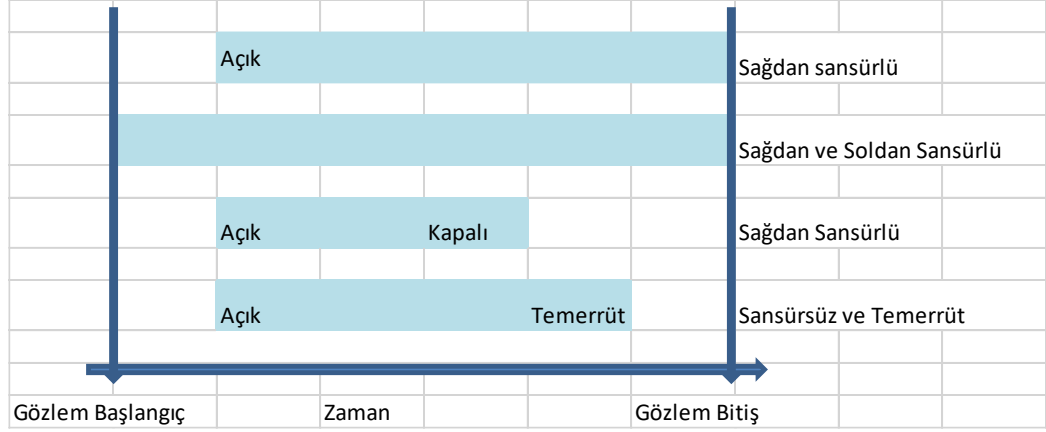
Ancak lojistik regresyon analizinin en büyük kısıtı büyük örneklem gerektirmesidir. Diğer varsayımlarının karşılanması nispeten kolaydır.

4.2 Cox Regresyon Analizi

Cox regresyon analizi, seçili bağımsız değişkenlerin bir olayın gerçekleşme süresi üzerindeki etkisinin modellenmesinde kullanılan bir yöntemdir. Cox regresyon analizi, yarı parametrik bir analiz türüdür ve bu analiz, birisi parametrik, diğeri ise parametrik olmayan iki temel bileşenden oluşmaktadır. Bu bileşenlerden β regresyon katsayısı, parametrik, $h_o(t)$ temel tehlike (hazard) fonksiyonu ise parametrik değildir.

Cox regresyon analizi, sağkalım analizi olarak adlandırılan tahmin modelleri içinde önemli bir yere sahiptir. Literatürde farklı sağkalım modelleme yöntemleri bulunmaktadır. Ancak Cox regresyon analizi bu yöntemler içinde en yaygın olarak kullanılan yöntemlerden birisidir. Kredi skorlama alanında en çok kullanılan sağkalım analiz yöntemi ise Cox Oransal Tehlike (Cox Proportional Hazard) modelidir. Bu model sayesinde kredi borçlusunun özelliklerini yansıtan değişkenler vasıtasıyla borçlunun ne zaman temerrüt edeceği tahmin edilir ve dolayısıyla temerrüt edip etmeyeceği de tahmin edilmiş olur.

Cox regresyon analizinde kritik konulardan birisi verinin sansürlü olmasıdır. Sansürleme (Kleinbaum ve Klein, 2005: 5), sağkalım süresi, hakkında kısmi bir bilgiye sahip olunup, kesin bir bilgiye sahip olunmadığı durumlarda oluşmaktadır. Cox regresyon analizi, sansürleme olarak ifade edilen bu sorunun üstesinden gelmek için geliştirilen bir yöntemdir. Diğer bir ifade ile bu analiz yöntemi sansürlü veri ile tahmin modeli kurmaya imkân tanımaktadır. Cox regresyon analizinde veri sansür çeşitleri Şekil 4.1'de görüldüğü gibidir.



Şekil 4.1. Cox regresyon analizi sansür çeşitleri

Araştırmanın başlangıç tarihinden sonra kullanılan ve araştırma bitiş tarihi itibarıyla canlı olan kredilerde temerrüt olayı gerçekleşmemiş ve bu durumda Şekil 4.1’de görülen sağdan sansür olayı meydana gelmiştir. Araştırmanın başlangıç tarihinden önce kullanılan ve araştırma bitiş tarihi itibarıyla canlı olan kredilerde, şekilde görüldüğü gibi sağdan ve soldan sansür meydana gelmiştir. Araştırmanın başlangıç tarihinden sonra kullanılan ve araştırma bitiş tarihi öncesinde kapanan kredilerde temerrüt olayı gerçekleşmemiş ve şekilde görüldüğü gibi sağdan sansür meydana gelmiştir. Son durumda ise kredi araştırmanın başlangıç tarihinden sonra kullanılmış ve temerrüt, araştırmanın bitiş tarihinden önce gerçekleşmiştir. Bu durumda sansür söz konusu değildir.

Cox regresyon analizinin matematiksel denklemleri (detaylı bir anlatım için bkz., Kleinbaum ve Klein, 2005: 94-96) aşağıda görüldüğü gibi açıklanabilir. Cox regresyonun ana denklemi,

$$h_i(t) = h_0(t)e^{(\beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik})} \quad (4.4)$$

olarak ifade edilir.

Bu formüle göre; $h_0(t)$, temel tehlike (hazard) fonksiyonudur. $e^{(\beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik})}$ ise üsteldir ve bağımsız değişkenlerin doğrusal

fonksiyonudur. Bu formüle göre t anındaki tehlike iki fonksiyonun çarpımıyla bulunur. Bu denklemdaki X'ler açıklayıcı değişkenlerdir. Bu değişkenlerin zamandan bağımsız olduğu varsayılmaktadır. Bu varsayım ihlal edildiğinde Cox regresyon analizinin temel denklemini bazı modifikasyonlar yapılarak kullanılabilir. Cox regresyon analizinin temel denklemini, zamana bağımlı değişkenleri içerecek şekilde dönüştürülebilir, bu dönüşüm uygulandıktan sonra Cox regresyon analizi, Genişletilmiş Cox regresyon analizi olarak ifade edilmektedir. Genişletilmiş Cox regresyon analizinin temel denklemini aşağıda görüldüğü gibidir.

$$h(t, X(t)) = h_0(t) \exp[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{i=1}^{p_2} \delta_i X_i(t)] \quad (4.5)$$

Yukarıdaki denkleminde görülen $X(t)$, hem zamana bağımlı değişkenleri hem de zamandan bağımsız değişkenleri temsil etmektedir.

4.2.1 Cox regresyon analizinin varsayımları

Bu bölümde Cox regresyon analizinin temel varsayımları (Sağkalım analizi bağlamında varsayımların değerlendirilmesi hususunda bkz., Tabachnick ve Fidel, 2015) anlatılmaya çalışılacaktır.

Birinci varsayım, belli bir zaman aralığı içinde yapılan sağkalım analizinde sağkalım koşulları zaman içinde değişime uğramamalıdır. Diğer bir ifade ile çalışmanın başlangıç döneminde etkili olan faktörlerin çalışma sonuna kadar etkili olmaya devam etmeleri gerekmektedir.

İkinci varsayım, Cox regresyon analizinin büyük örneklem ile yapılması gerekmektedir. Çünkü bu analizde parametre tahmininde en çok olabilirlik yöntemi kullanılmakta ve bu yüzden daha güvenilir tahminler yapılabilmesi için büyük örneklem ile çalışılması gerekmektedir.

Üçüncü varsayım, Cox regresyon analizinde tehlike (hazard) fonksiyonları birbirlerine oransal olmalı, diğer bir ifadeyle, bir değişkenin etkisi tüm zamansal noktalarda aynı olmalıdır. Bu varsayım, Cox regresyon analizinin temel varsayımıdır. Oransal tehlike varsayımının testi için grafiksel yöntemler ve sayısal testler kullanılmaktadır. Tabachnick ve Fidel (2015: 539-40) sadece grafiksel yöntemlerle yetinilmemesini ve istatistiksel testlerin de yapılması gerektiğini belirtmektedirler. Zamanla etkileşim değişkenleri türetilerek, bu değişkenlerin anlamlılığı test edilir ve zamanla etkileşim değişkenleri anlamlı bulunmayıp modele alınmazsa, oransallık varsayımının sağlandığı kabul edilmektedir.

Dördüncü varsayım, sansürlü gözlemler ile sansürlü gözlemler birbirinden çok farklılaşmaması gerekmektedir.

Beşinci varsayım, parametre kestirimlerinde standart hatanın artmasına neden olan çoklu doğrusallık olmaması gerekmektedir.

4.2.2 Cox regresyon analizinin avantajları ve kısıtları

Lojistik regresyon analizi ile sadece müşterinin temerrüt edip etmeyeceği tahmin edilirken, Cox regresyon analizi, müşterinin temerrüt edip etmeyeceğinin tahminine ilaveten ne zaman temerrüt edeceğini de tahmin etmeye imkân tanır. Bu özelliği Cox regresyon analizinin en önemli avantajlarından biri olmaktadır.

TFRS 9'a, "beklenen kredi zararlarının ölçüleceği azami süre[nin], işletmenin kredi riskine maruz kaldığı azami sözleşme süresi" olduğunu belirtmektedir. Standarda göre; "ömür boyu beklenen kredi zararları açısından işletme, finansal aracın beklenen ömrü boyunca temerrüde düşülmesi riskini tahmin" etmek zorundadır. Örneğin lojistik regresyon analizinde genellikle gözlem süresi 12 ay olarak belirlenmekte ve dolayısıyla 12 aylık tahminler yapılmaktadır. Cox regresyon analizi ile 12 ay gibi bir süre ile sınırlı kalmadan kredi ömrünü kapsayacak bir süre için tahmin yapılabilir. Lojistik regresyon analizi ile belli bir

zaman dilimi içinde müşterinin temerrüt edip etmeyeceği tahmin edilir. Diğer bir ifade ile bir noktadan bakılıp, o noktadan itibaren sonraki 12 ayda müşterinin temerrüt edip etmeyeceği modellenir. Cox regresyon analizinde ise bir noktadan bakılıp sonrası modellenmez. Aksine belirlenen bir süreç içinde yani araştırmanın başlangıç ve bitiş tarihleri arasında müşterinin ne zaman temerrüt edeceği modellenir.

Cox regresyon analizinde herhangi bir dağılımsal varsayım gerekmemektedir. George vd. (2014)'nin vurguladığı üzere Cox regresyon analizinin önem kazanmasının sebebi, araştırmacıların kullandıkları veride özgün bir sağkalım dağılımının olduğunu varsaymak zorunda olmamalarıdır. Dağılımla ilgili varsayım gerektiren istatistiksel analizler kullanıldığında sıklıkla karşılaşılan sorunlardan birisi, verinin dağılım şartını sağlamamasıdır. Araştırmacılar bu durumda ilave zaman ve emek harcayarak veri dönüşümü yapmaktadırlar. Veri dönüşümü sayesinde dağılım şartı sağlanmakta ama verinin içeriğinde olan bazı bilgiler de kaybedilmiş olabilmektedir. Cox regresyon analizinde ise verinin dağılım şartı aranmadığından, Cox regresyon analizi ilave zaman ve emek maliyetine sebep olmamasının yanı sıra orijinal veri ile çalışmaya izin vermektedir. Bu nedenle Cox regresyon analizi uygulamalarda oldukça kullanışlı bir analiz yöntemidir.

5. UYGULAMA

Bu bölümde lojistik regresyon analizi ile Genişletilmiş Cox regresyon analizi uygulaması aşama aşama anlatılacaktır.

5.1 Uygulamanın Amacı ve Hipotezi

Uygulamanın temel amacı, bankaların, TFRS 9'un getirdiği yasal zorunluluklardan birisini karşılayabilmesi hususunda alternatif bir çözüm önerisi sunmaktır. Uygulamada, Cox regresyon ve lojistik regresyon analizleri ile bankaların kurumsal (tüzel) kredi müşterilerine yönelik TFRS 9 uyumlu bir temerrüt olasılığı modeli geliştirilecektir. Esas olarak, derlenen veriye Cox regresyon ve lojistik regresyon analizleri uygulanarak, sonuçlar karşılaştırılacak ve hangi analizle daha kesin sonuçlar elde edildiğinden hareketle temerrüt olasılığına dair bir tahmin modeli önerilecektir.

Bankalar, genellikle lojistik regresyon analizini kullanmayı tercih etseler de literatüre ve yönetime dair bölümlerde değinilen bazı akademik çalışmalarda görüldüğü üzere Cox regresyon analizi, lojistik regresyon analizine bir alternatif olarak gündemdedir. Cox regresyon analizinin bir alternatif olup olmadığını bankacılık verileriyle test etmek amacıyla 'Cox regresyon analizi, lojistik regresyon analizine göre daha açıklayıcı ve kesin (robust) sonuçlar üretmektedir' hipotezi geliştirilmiştir.

5.2 Uygulamadaki Modelleme Yaklaşımı

Çoğu banka kurumsal segmentteki kredi müşterilerini alt-segmentlere ayırmakta ve pazarlama, tahsis, izleme ve takip faaliyetlerini bu alt-segmentler bazında yürütmektedir. Bu yüzden bu alt-segmentler bazında modellere ihtiyaç vardır. Geliştirilen modellerin banka ihtiyaçlarını dikkate alan kullanılabilir modeller olması gerekmektedir. Bu çalışmada bir bankanın gerçek verileri kullanılarak bir

yıllık süre zarfında müşterilerin temerrüt edip etmeme ihtimallerini hesaplamaya yönelik iki tane regresyon modeli geliştirilecektir.

5.2.1 Modellerde kullanılan temerrüt tanımı

Temerrüt tanımı tüm araştırmayı etkileyeceği için net bir şekilde tanımlanmalıdır. Müşterinin kesit tarihinde riski devam eden projelerinde veya performans dönemi içinde yeni kullandığı projelerinden herhangi bir tanesinde, ödemesini 90 günden fazla geciktirirse veya geç gün sayısı 90 gün olmadan donuk alacak olarak sınıflandırılıp takip hesaplarına virmanlanırsa veya iflas veya iflas erteleme ile ilgili bir karar alınırsa, müşteri temerrüt etmiş olarak kabul edilir. Temerrüt tarihi olarak hangisi önce gerçekleşmişse o kabul edilecektir.

5.2.2 Modellerin kurulacağı dönem

Modeller için, 2011 tarihinden başlayıp 31 Aralık 2015 tarihine kadar süren 6 aylık periyotlar ile 10 kesit tarihi için veri oluşturulmuştur. Kesit tarihleri sırasıyla şöyledir: 30/06/2011, 31/12/2011, 30/06/2012, 31/12/2012, 30/06/2013, 31/12/2013, 30/06/2014, 31/12/2014, 30/06/2015 ve 31/12/2015. 5 yıllık bir veri seti ile çalışılacağı için kurulan modelin banka müşterilerinin genel yapısını yansıtmaya imkânı daha fazladır.

Müşterilerin, her bir kesit tarihinden sonraki bir yıl içindeki performansları değerlendirilecektir. Yani, 10 ayrı performans dönemi olacaktır. Bu kesitler birleştirilerek analizler yapılacaktır.

5.2.3 Modellerde yer alan müşteriler

Veri seti banka tarafından maskelenmiş bir şekilde paylaşılmıştır. Banka tarafından her bir müşteri için rassal bir şekilde yeni bir tekil ID oluşturulmuştur. Böylece firmaların gerçekte hangi firma olduğu gizlenmiş ve

müşteri gizliliği korunmuştur. Çalışmalar bu maskelenmiş veri üzerinden yapılmıştır.

Kesit tarihi itibariyle kredi riski olan, gecikmesi olmayan, birinci grup olarak sınıflandırılan ve 6 aydan fazla süredir banka ile çalışıyor olan (yani eski müşteri) genel kurumsal portföyün bir alt-segmentinde yer alan müşterileri için modelleme yapılmıştır. Bu alt-segment, yıllık geliri 3 ile 30 milyon TL arasında olan veya bankanın 1 ile 10 milyon TL arası limit açtığı müşterilerden oluşur. Diğer taraftan, kesit tarihi itibariyle 1 gün bile gecikmesi olan müşteriler modelleme kapsamına alınmamıştır. Bu hususta kesit tarihinden önceki 6 veya 12 ayda, belli bir günden fazla –örneğin 60 gün– gecikme yaşanmışsa modellemeden çıkarılması gibi farklı yaklaşımlar da söz konusudur. Bu hususta daha farklı yaklaşımlar da vardır. Örneğin Sıddıqi (2006: 3) son 24 ayda 60 gün üzeri gecikmenin ikiden fazla yaşandığı müşterilerin modelleme kapsamından çıkarılabileceğini belirtmektedir. Gecikmesi olan müşterilerin temerrüt etme veya yapılandırmaya konu olma ihtimali, gecikmesi olmayanlara göre daha yüksektir. Gecikmesi olan ve olmayan müşteriler aynı anda modellenmeye dahil edildiğinde homojen bir yapı ile çalışılmamaktadır. Ayrıca, bir müşteri kredi ödemesini zamanında yapmayarak ciddi bir erken uyarı sinyali vermiş olmaktadır.

Birinci grup (standart nitelikli kredi) ibaresini biraz daha detaylandırmak modellemeye konu kitlenin anlaşılmasını daha kolaylaştıracaktır. Birinci grup krediler ve ikinci grup krediler, canlı kredi olarak sınıflandırılmaktadır. Birinci grup (standart nitelikli krediler) ve ikinci grup (yakın izlemedeki krediler) kavramları, Karşılıklar Yönetmeliği'nde bulunan kavramlardır. Bu çalışmada Karşılıklar Yönetmeliği'nde tanımlandığı şekliyle kullanılmışlardır. BDDK (2018)'nin internet sitesi üzerinden güncel Karşılıklar Yönetmeliği'ne erişilip tanımların detaylarına bakılabilir. TBB (2018)'nin internet sitesi üzerinden Karşılıklar Yönetmeliği'nin eski versiyonlarına ulaşılabilir. Modelleme çalışmasının yapıldığı yıllarda Karşılıklar Yönetmeliği'nin eski versiyonu yürürlükte olduğundan uygulamada eski versiyon kullanılmıştır.

İkinci grup (yakın izlemedeki krediler) krediler, bu çalışmada kapsam dışında bırakılmışlardır. Şöyle ki bazı müşteriler gecikmesi olmasa da yapılandırma gibi başka sebeplerden dolayı ikinci grup olarak sınıflandırılabilir. Bir kredinin ikinci grupta sınıflandırılması kredi riskinde zaten önemli bir artış olduğunu göstermektedir. Homojen bir yapı tesis edebilmek için ikinci grup krediler çalışmanın kapsamı dışında tutulmuştur.

Bankalar zaten gecikme yaşandığında veya müşteri ikinci grup olarak sınıflandırıldığında çeşitli aksiyonlar almaktadırlar. Geliştirilecek modeller ile gecikme yaşanmadan daha erken bir dönemde müşteriler hakkında tahmin modelleri kurulması hedeflenmektedir.

Geliştirilecek olan model davranışsal bir model olacağı için müşterilerin bankada belli bir geçmişinin olması gerekmektedir. Piyasa uygulamaları doğrultusunda, banka ile 6 aydır çalışan müşteriler eski müşteri olarak kabul edilmiştir. Çalışmada, kesit tarihinde açık olan (riski devam eden) nakdi veya gayri nakdi projelerinden herhangi birisinin kullandırım tarihi 6 aydan eski olan müşteri eski müşteri olarak kabul edilmektedir. Ayrıca, kesit tarihi itibarıyla riski olan projelerin hepsi yeni -yani son 6 ay içinde yapılmış- olsa bile müşterinin son 12 ayda kapanan herhangi bir nakdi projesi varsa eski müşteri olarak kabul edilmiştir. Sürekli 3-4 aylık kısa vadeli kredi kullanan müşterilerin 6 ay kriterinden dolayı kapsam dışında kalmasının önüne geçilebilmesi için bu ilave kriter eklenmiştir.

Modellemede yer alan birinci grup olarak sınıflandırılan eski müşterilere ilişkin gözlem setinin elemeler öncesi dağılımı Tablo 5.1’de görüldüğü gibidir.

Tablo 5.1. Elemeler öncesi gözlem setinin dağılımı

Gözlem Sayısı	Temerrüt Etmeyen (Canlı) Müşteri Sayısı	Temerrüt Eden Müşteri Sayısı	Temerrüt Eden Müşteri Yüzdesi
98.053	95.979	2.074	% 2,1

Tablo 5.1’de görüldüğü gibi temerrüt eden müşteri yüzdesi gözlem setinin % 2,1’dir.

5.2.4 Modellere ilişkin örneklemin oluşturulması

Yukarıda verilen gözlemlerden oluşan veri seti, kesit bazlı oluşturulduğu için modellemeye başlanmadan önce çoklama sorununun giderilmesi amacıyla bazı elemelerin yapılması gerekmektedir. Çünkü bir kesit tarihi baz alınıp sonraki 12 ayda temerrüt olayının gerçekleşip gerçekleşmediğine bakıldığı için, aynı olay birden fazla kesit tarihine kaydedilebilir. Örneğin müşteri Kasım 2014’te temerrüt etmiş ise, 31 Aralık 2013 kesit tarihinde sonraki 12 ayda temerrüt etmiş mi diye baktığımızda temerrüt etmiş olarak kaydedeceğiz. 30/06/2014 kesit tarihinde de aynı olay nedeniyle müşteriye temerrüt olarak kaydetmiş olacağız. Bu örnekte görüldüğü üzere aynı temerrüt olayı iki defa kullanıldığı için bir çoklama sorunu mevcuttur.

Bu çoklama sorununu tespit etmek için öncelikle her bir kesit tarihi için performans döneminde (yani sonraki 12 ay) hangi olay gerçekleşmişse onu gösteren bir etiket tanımlanması gerekmektedir. Çünkü bir müşteri sonraki 12 ayda ya temerrüt eder veya etmeyebilir. Çalışmada çoklamaya konu temerrütlerin elenmesi hususunda; müşterinin temerrüt etiketi ile işaretlendiği kesit tarihleri belirlenmiş, aynı temerrüt olayından dolayı birden fazla kesit tarihinde işaretleme yapıldığı durumlarda ilgili en küçük kesit tarihi alınmış ve diğerleri silinmiştir.

Bu işlemlerden sonra çoklamaya konu canlı kredilerin elenmesi hususunda; müşterilerin canlı olarak etiketlendiği kesit tarihleri belirlenmiş, müşteriler kesit tarihlerine göre sıralanmışlar ve yeni bir bilgi içermeyen kesit tarihleri silinmiştir. Bir örnekle anlatmak gerekirse, bir müşteri tüm kesit tarihlerinde canlı olarak etiketlenmiş ise, örneğin 30/06/2011 kesit tarihinde sonraki 12 aya baktığımızda 30/06/2012'e kadar bakmış oluruz. Diğer taraftan, 30/06/2012'de baktığımızda ise 30/06/2013'e kadar olan döneme bakmış oluruz. Bu durumda, örneğin 31/12/2011 kesit tarihine bakmaya gerek kalmamaktadır. O dönem çıkarıldığında müşterinin sonraki performansı hakkında bilgi kaybımız olmamaktadır.

Bu 2 aşamalı eleme işleminden sonra geri kalan canlı kavramı ile isimlendirilen temerrüt etmeyen müşteriler ile temerrüt eden müşteriler alınarak model kurulmuştur. Elemeler yapıldıktan sonra gözlem setinin dağılımı Tablo 5.2'de görüldüğü gibidir.

Tablo 5.2. Elemeler sonrası gözlem setinin dağılımı

Gözlem Sayısı	Temerrüt Etmeyen (Canlı) Müşteri Sayısı	Temerrüt Eden Müşteri Sayısı	Temerrüt Eden Müşteri Yüzdesi
57.174	55.492	1.682	% 2,9

Tablo 5.2'de görüldüğü gibi temerrüt eden müşteri yüzdesi artarak % 2,9 değerine ulaşmıştır.

5.2.5 Modellerdeki değişkenler

Modellerdeki bağımlı ve bağımsız değişkenler çalışmamızın bundan sonraki bölümünde ayrı ayrı anlatılacaktır.

5.2.5.1 Bağımlı değişken

Bağımlı değişken; sonraki bir yıl içinde temerrüt etmiş ve sonraki bir yılda geri ödemelerini yapmış yani temerrüt etmemiş olmak üzere 2 kategoriden oluşmaktadır. Temerrüt etmeme durumu daha önceki açıklamalarda değinildiği gibi 'canlı' kavramı ile de ifade edilebilecektir.

5.2.5.2 Bağımsız değişkenler

Bağımsız değişken seçimi sürecinde Sıddıqı'nın (2017: 133-140) yaklaşımı uygulanmıştır. Sıddıqı'nın yaklaşımını özetlemek gerekirse, teknik olarak binlerce değişkenle model kurmaya başlamak mümkündür. Çoğu banka böyle büyük veri setlerine sahiptir. Fakat çok fazla değişkenle çalışmak çok etkin bir yol değildir. Çünkü bu durumda modelleme sürecinin kontrollü bir şekilde sürdürülmesi zordur ve kontrolden çıkabilir. Bu nedenle, öncelikle bankanın sağlayabildiği değişkenlerin tam listesi hazırlanmakta, daha sonra modelleme ekibi, validasyon ekibi, risk yönetimi vb. ilgili tarafların temsilcileriyle bu değişkenlerin hangilerinin modelleme sürecine dâhil edilip edilmeyeceği hususunda müzakere edilmektedir. Bu müzakereler sonucunda sınırlı sayıda değişken tespit edilebilmektedir. Değişkenlerin tespit işleminde Sıddıqı'nın dikkate aldığı birçok kriter vardır. Bunlardan en önemlisi, değişkenlerin beklenen açıklayıcılık gücüdür. Banka genelde bir modelleme geçmişine sahip olduğu için ilgili analistler hangi değişkenlerin açıklayıcılık gücünün yüksek olduğu konusunda bir tecrübeye sahiptirler. Ayrıca, literatürdeki yayınlar incelendiğinde 30 ile 50 arasında değişkenin açıklayıcılık gücünün genelde yüksek çıktığı görülmektedir. Diğer taraftan, modelde kullanılacak verinin gelecekte de elde edilebiliyor olması gerekmektedir. Gelecekte yasal nedenlerden dolayı elde edilemeyecek bir verinin kullanılmaması gerekir.

Im vd. (2012) çok büyük bir veri seti kullanarak yaptıkları çalışmada değişkenleri tespit ederken Sıddıqı ile benzer bir yöntem kullanmışlardır. Veriyi sağlayan finansal kuruluş binlerce değişkeni içeren büyük bir veri ambarına

sahiptir. Bu deęişkenler arasından ilgili saha uzmanları tarafından potansiyel 75 deęişken belirlenmiş ve modelleme çalışmasına bu 75 deęişkenle devam edilmiştir.

Çalışmamızda kullanılan veriyi sağlayan banka, davranış ve başvuru modellerine yönelik kapsamlı bir veri ambarına sahiptir. Bu veri ambarından Sıddıqı'nın yaklaşımı doğrultusunda davranışsal temerrüt olasılığı modellemesine uygun, doluluk oranı yüksek, iş mantığı açısından kabul edilebilir, birbirleriyle düşük korelasyona sahip olan ve ön testlerden geçen 28 deęişken belirlenmiştir. Bu deęişken belirleme sürecinde ilgili tüm bağımsız deęişkenlerin modele dâhil edilmesi ve ilgisiz deęişkenlerin elenmesi sağlanmıştır.

Çalışmadaki modellerimizde kullanılan bağımsız deęişkenler, Ek A'da görülmektedir.

5.2.6 Modellerde kullanılan verilerin analizi

Sıddıqı (2017: 21)'nin belirttiği üzere bir regresyon modelinin kalitesini belirleyen en önemli unsur veri kalitesidir. Veri kalitesi hususunda banka tarafından yoğun çalışmalar yapılmıştır. Bu çalışmada kullanılan veri; ekransal kontroller (veri setinde raporlanan bilgi ile ilgili ekrandaki bilginin aynı olup olmadığının teyidi) ve çapraz sorgu kontrolleri (bazı mantıksal kurallar çerçevesinde verinin teyidi, örneğin son 3 ayda gecikmeli ödenen taksit sayısı son 6 ayda gecikmeli ödenen taksit sayısından büyük olamaz vb.) gibi kontrol aşamalarından geçmiş ve doğruluğu teyit edilmiş bir veridir. Bu yüzden verinin doğruluğu hususunda ilave bir çalışma yapılmamıştır.

Ancak, veride karşılaşılan sorunlardan bir tanesi, sıfır frekans ve tam ayırma sorunudur. Sıfır frekans, bağımsız deęişkenlerin bazı kategorileri için gözlem olmamasıdır. Tam ayırma ise bağımlı deęişkenin aldığı deęerin bağımsız deęişkenin bir kategorisi ile % 100 ilişkili olmasıdır. Örneğin temerrüt eden bütün müşterilerin olumsuz istihbarat kaydı varsa istihbarat kaydı deęişkeni

modele alındığında tam ayırma olur. Hem sıfır frekans hem de tam ayırma durumlarında “regresyon katsayıları ve güven aralıkları aşırı büyük kestirilebilir.” (Alpar, 2017: 613-4).

Sıfır frekans sorununa çözüm olarak, örneklem büyüklüğünü arttırarak ilgili değerler için yeterli gözlem çalışmaya ilave edilmeli veya değerleri birleştirme yoluna gidilmelidir. Çalışmada sürekli değişkenler yeniden sınıflandırılarak bu sorun giderilmiştir. Şöyle ki oluşturulan yeni değişkenin her kategorisinin yeteri kadar temerrüt etmiş gözlem içermesi sağlanmış ve böylece sıfır frekans sorunu yaşanmamıştır.

5.2.7 Modelleri değerlendirme analizleri

Kurulan modellerin doğruluğunu değerlendirmek için farklı yöntemler olsa da bu yöntemler içinde en önemlisi ROC eğrisi ve GINI katsayısı analizidir.

5.2.7.1 Roc eğrisi ve gini katsayısı analizi

ROC eğrisi analizi, model örnekleminde canlılar ve temerrütlerin dağılımının eşit olmadığı durumlarda, iki tanılama grubu arasındaki ayırt ediciliği ölçmek için kullanılmaktadır. ROC değerinin 0 olması ‘Canlılar’ ve ‘Temerrütler’ arasında hiçbir ayırım gücü olmadığını, ROC değerinin 100 olması ise; ayırım gücünün mükemmel olduğunu göstermektedir.

ROC eğrisi ile GINI katsayısı arasında bir ilişki bulunmakta ve bu ilişki;

$$GINI = 2 \times ROC - 100$$

olarak ifade edilmektedir.

GINI katsayısı genellikle 0 ila 100 arası değer olur. GINI katsayısı 0 olduğunda, modelde canlı ve temerrüt eden müşteriler arasında hiçbir ayırım gücünün

olmadığını, GINI katsayısının 100 olması ise, model performansının mükemmel ayırım gücüne sahip olduğu anlamına gelmektedir.

5.2.7.2 Hata matrisi analizi

Hata Matrisi (Confusion Matrix) analizi, modelin tahminleri ile gerçekleştirmelerin karşılaştırıldığı ve modelin doğru tahmin yüzdesini gösteren bir analizdir. Burada değerlendirme yapılırken Birinci Tip Hata ve İkinci Tip Hataya bakılır. Birinci Tip Hata, gerçekte temerrüt eden ama modelde yanlışlıkla canlı olarak tanımlanan firmalarla ilgidir. İkinci Tip Hata, gerçekte temerrüt etmeyen ama modelde yanlışlıkla temerrüt olarak tanımlanan firmalarla ilgidir. Birinci ve ikinci tip hata değerlerinin hesaplanmasında eşik değerin diğer bir ifade ile kesim noktasının önemli bir yeri vardır.

Hata matris analizinde önemli olan kesim noktasının belirlenmesidir. Kesim noktasının (cut-off) belirlenmesi hususunda farklı yaklaşımlar vardır. Standart uygulama olarak kesim noktası 0,5'tir. Belirlenen kesim noktasına göre doğru sınıflandırılan temerrüt gözlemlerinin sayısı değişecektir. Modina ve Pietrovito (2014: 1550), örneklemin temerrüt oranını da içerek şekilde 0,03, 0,05, 0,10 ve 0,15 gibi farklı kesim değerlerine göre model kurulmasını ve bu modellerin sonuçlarının karşılaştırılarak kesim noktasının belirlenmesini tavsiye etmektedir. Diğer bir alternatif, optimal kesim noktasının tespit edilip kullanılmasıdır. Bu hususta Bottazzi vd. (2011: 394) optimal yani birinci tip hata ve ikinci tip hataları minimize eden bir çözüm sunmuştur.

Uygulama regresyon modelleri kurulurken uygulanan yaklaşımlardan birisi, canlı müşteri adedi kadar temerrüt eden müşteri ile model kurmaktır. Temerrüt ve canlı müşteri adedinin dengeli olduğu bu durumlarda kesim noktası 0,50 olarak alınmaktadır. Bu çalışmada kullanılan yaklaşım, popülasyonun temerrüt oranını aynen korumak yönündedir. Çünkü kullandığımız veri setinde canlı müşteri ile temerrüt eden müşteri adetleri arasında dengesizlik mevcuttur.

5.3 Lojistik Regresyon Analizi Uygulaması

(4.3) nolu lojistik regresyon eşitliğinin parametrelerini belirlemek amacıyla lojistik regresyon analizi uygulanmıştır. Lojistik regresyon analiz uygulaması SAS Enterprise Guide 7.1 ve SAS Enterprise Miner Client 4.1 ile yapılmıştır. Veri, Guide üzerinde modellemeye hazır hale getirilmiştir. Ardından Miner ile modelleme yapılmıştır. Lojistik regresyon analizinin uygulama süreci veri setinin SAS Miner programına aktarılması ile başlamaktadır. 'StatExplore' düğümüyle betimleyici istatistiksel analizler yapıldıktan sonra, Veri Bölme düğümüyle veri seti ikiye ayrılmakta, ardından Etkileşimli Gruplama düğümüyle değişkenlere WoE dönüşümü uygulanmakta ve son olarak Skorkart ve Regresyon düğümleriyle model geliştirilmektedir. Model geliştirildikten sonra Kesim Noktası düğümüyle hata matrisi analizi yapılmıştır.

5.3.1 Değişken analizleri

Tek değişkenli analizin amacı, bir müşterinin temerrüt etme olasılığı ile ilişkili uygun değişkenleri belirlemektir. Tek değişkenli analizle ayırım gücü yüksek olan değişkenler tespit edilmektedir. Canlı/temerrüt ayırımını tahmin etmek için mükemmel güce sahip değişkenler, bu adımda belirlenmektedir. GINI katsayısı ya da Bilgi Değeri gibi istatistiksel analizler, değişkenlerin anlamlılık gücünü ölçmek için kullanılmaktadır. Bu safhada değişkenler iş mantığı açısından da kontrol edilmektedir. Her bir değişken iş mantığına göre anlamlı olmalıdır. İş mantığı açısından uygun olmayan değişkenler elenmektedir.

Çalışmada kullanılan değişkenler, tek değişkenli analizlere tabi tutulmuş ve bu açıdan anlamlı olan değişkenlerle modellemeye başlanmıştır.

Korelasyon analizi, iki değişken arasındaki ilişkinin ilişki gücünü analiz etmek için kullanılan bir istatistiksel ölçüm yöntemidir. Tahmin modeli geliştirme sürecinde, değişkenlerin herhangi ikisi arasında yüksek korelasyon olmamalıdır.

Bu nedenle, tahmin modeli geliřtirmede deęiřken seęimi, tek deęiřkenli analiz sonrasında korelasyon analizine baęlıdır.

Korelasyon katsayısı -1 ila 1 arası deęerler alır. Model geliřtirilirken aralarında yüksek korelasyon olan deęiřkenlerden açıklayıcılık gücü düşük olanlar elenmiřtir. Pozitif veya negatif yönde -0,5 ile 0,5 arası deęere sahip deęiřkenler modelleme kapsamına alınmiřtır.

Yapılan korelasyon analizi sonucunda, etkileřimli grupta öncesi deęiřkenler arasındaki maksimum pozitif korelasyon RM2 ile RM5 deęiřkenleri arasındadır ve korelasyon katsayısı deęeri 0,436'dır. Etkileřimli grupta öncesi deęiřkenler arasındaki maksimum negatif korelasyon TEMINAT ile CEK1 deęiřkenleri arasındadır ve korelasyon katsayısı deęeri -0,479'dur.

5.3.2 Etkileřimli grupta

Etkileřimli Grupta düęümüyle deęiřkenlere monotonik bir yapı kazandırılır. řöyle ki bir deęiřken her bir sınıftaki temerrüt oranlarına göre azalan veya artan řekilde yeniden sınıflandırılır. Örneęin sürekli bir deęiřken belli aralıklarla 5 sınıfı ayrılır ve birinci sınıftan sonuncu sınıfa doęru temerrüt oranı ya artar ya da azalır.

Etkileřimli Grupta düęümü, seęilen kriterlere göre her bir deęiřkeni yeniden sınıflandırmakta ve her bir grup için WoE deęeri hesaplamaktadır. Deęiřkenlerin modellemede kullanılmadan önce tek tek incelenmesi ve daha monoton bir yapı tesis edilmeye çalıřılması gerekir. Deęiřkenler iř bilgisine sahip uzmanlar tarafından analiz edilerek iř tecrübesi grupta yansıtıldıęında deęiřkenlerin ayırım güçleri artabilmektedir.

GECIKME1 deęiřkeninin sistem tarafından hesaplanan WoE grupta Tablo 5.3'te görüldüęü gibidir.

Tablo 5.3. WoE değerleri

Değer	Grup	Temerrüt Adedi	Canlı Adedi	Toplam	Temerrüt Oranı	WoE
0 <= GECIKME1 < 1	1	445	29107	29552	0,015	0,6841
1 <= GECIKME1 < 3	1	127	3238	3365	0,038	-0,25804
3 <= GECIKME1 < 7	2	152	2791	2943	0,052	-0,58628
7 <= GECIKME1 < 16	2	189	1943	2132	0,089	-1,16632
GECIKME1 >= 16	2	264	1764	2028	0,13	-1,59717

Tablo 5.3'te görüldüğü üzere, her bir grup bazında temerrüt oranı artmış ve değişkene monoton bir yapı kazandırılmıştır.

Yukarıdaki açıklamalarda değinilen WoE (Weight of Evidence – Kanıt Ağırlığı) analizi, temerrüt oranlarına bağlı olarak değişkenlerin ayırt ediciliğinin ölçülmesidir. WoE, bağımsız değişkenlerin bağımlı değişkeni tahmin gücü hakkında bilgi verir. WoE değeri her bir grup için aşağıdaki formülle hesaplanır:

$$WoE_{grup} = \ln \left(\frac{\text{Gruptaki Canlı Müşteri Adedi} / \text{Toplam Canlı Müşteri Adedi}}{\text{Gruptaki Temerrüt Müşteri Adedi} / \text{Toplam Temerrüt Müşteri Adedi}} \right)$$

Modellemede değişkenlerin ham hali değil de etkileşimli gruplama düğümüyle üretilen WoE değerleri kullanılmaktadır.

Bilgi Değeri (Information Value) bulunurken gruptaki canlı müşteri yüzdesinden gruptaki temerrüt müşteri yüzdesi çıkarılır ve sonuç grubun WoE değeri ile çarpılır. Hesaplama formülü şöyledir:

$$IV_{grup} = [(\text{Gruptaki Canlı Müşteri Adedi} / \text{Toplam Canlı Müşteri Adedi}) - (\text{Gruptaki Temerrüt Müşteri Adedi} / \text{Toplam Temerrüt Müşteri Adedi})] * WoE_{grup}$$

Sıddıqi (2017: 185)'nin belirttiđi üzere, Bilgi Deđeri 0,02'den küçük ise tahmin gücünün olmadığını, 0,02 ile 0,1 arasında ise zayıf tahmin gücünün olduğunu ve 0,1 ile 0,3 arasına ise kuvvetli tahmin gücünün olduğunu gösterir.

Modellemedeki ana hedef, mümkün olan en az deđişkenle sağlam ve tutarlı tahminler yapmaktır. Açıklayıcılık gücü yüksek olmayan deđişkenleri elemek için çeşitli eşik deđerler belirlenmektedir. Bu çalışmada GINI'si 20'den küçük olan veya Bilgi Deđerleri 0,14'ten küçük olan deđişkenler elenerek modellemeye alınmamıştır.

Her bir deđişken için hesaplanan GINI katsayısı ve Bilgi Deđerleri Tablo 5.4'te görüldüğü gibidir. GINI ve Bilgi Deđerleri için belirlenen eşik deđerlerden dolayı 10 deđişken elenerek modelleme dışında bırakılmıştır. Elenen deđişkenler; CEK3, YAS2, LIMIT1, MEVDUAT, YAS1, LIMIT2, RM7, RM3, ISTIHBARAT ve GORUSME'dir.

Tablo 5.4. Lojistik modeldeki deęişkenlerin gini katsayıları ve bilgi deęerleri

Sıra	Deęişken	Hesaplanan Rol	GINI Katsayısı	Bilgi Deęeri
1	GECIKME1	Kabul Edildi	41,39	0,701
2	RM1	Kabul Edildi	33	0,511
3	PROJE	Kabul Edildi	37,5	0,505
4	RM8	Kabul Edildi	36,49	0,475
5	LIMIT3	Kabul Edildi	37,2	0,471
6	RM2	Kabul Edildi	25,77	0,377
7	RM6	Kabul Edildi	31,75	0,368
8	RM5	Kabul Edildi	24,27	0,368
9	RM10	Kabul Edildi	26,28	0,314
10	RM4	Kabul Edildi	29,34	0,312
11	TEMINAT	Kabul Edildi	29,61	0,303
12	GECIKME2	Kabul Edildi	20,03	0,286
13	ORTAK	Kabul Edildi	27,66	0,271
14	FINANSAL2	Kabul Edildi	26,61	0,249
15	CEK2	Kabul Edildi	22,07	0,233
16	CEK1	Kabul Edildi	21,87	0,215
17	RM9	Kabul Edildi	21,23	0,196
18	FINANSAL1	Kabul Edildi	23,37	0,19
19	CEK3	Red Edildi	12,46	0,12
20	YAS2	Red Edildi	15,78	0,087
21	LIMIT1	Red Edildi	13,85	0,079
22	MEVDUAT	Red Edildi	11,48	0,045
23	YAS1	Red Edildi	9,775	0,032
24	LIMIT2	Red Edildi	7,526	0,025
25	RM7	Red Edildi	7,448	0,024
26	RM3	Red Edildi	6,271	0,018
27	ISTIHBARAT	Red Edildi	4,213	0,014
28	GORUSME	Red Edildi	0	0

5.3.3 Öğrenme ve doğrulama veri setleri

Uygulamalarda, genel olarak, tüm örneklem ile model geliştirmek yerine öğrenme (training) ve doğrulama (validasyon) veri setleri oluşturularak model kurulmaktadır. Genelde gerçek veri setinin % 70'i öğrenme, % 30'u da doğrulama için ayrılmaktadır. Çalışmada bu işlem Veri bölme düğümüyle yapılmıştır. Yapılan işlem öncesi değerler aşağıdaki Tablo 5.5'te görüldüğü gibidir.

Tablo 5.5. Gerçek veri seti

Gerçek Veri Seti		
Bağımlı Değişken Değeri	Gözlem Adedi	Gözlem Yüzdesi
Canlı	55.492	97,1%
Temerrüt	1.682	2,9%
Toplam	57.174	100,0%

Aşağıdaki Tablo 5.6'da görülen bölme işlemi iki yöntemle yapılmaktadır. Bunlar; basit rassal bölme ve tabakalı rassal bölmedir. Burada tahmin edilecek değişken ikili yapıda olduğu için tabakalı rassal bölme yöntemi kullanılmıştır. Böylece temerrüt etmiş ve canlı müşterilerin oranı, ana veri setinde, öğrenme veri setinde ve doğrulama veri setinde birbirine yakın değerlerdedir. Bu değerler Tablo 5.6'da görüldüğü gibidir.

Tablo 5.6. Öğrenme ve doğrulama veri setleri

Öğrenme Veri Seti		
Bağımlı Değişken Değeri	Gözlem Adedi	Gözlem Yüzdesi
Canlı	38.843	97,1%
Temerrüt	1.177	2,9%
Toplam	40.020	100,0%

Doğrulama Veri Seti		
Bağımlı Değişken Değeri	Gözlem Adedi	Gözlem Yüzdesi
Canlı	16.649	97,1%
Temerrüt	505	2,9%
Toplam	17.154	100,0%

Tablo 5.6 incelendiğinde oranların birbirine oldukça yakın olduğu görülmektedir.

5.3.4 Değişken seçimi ve regresyon

Yapılan araştırmalarda birden fazla değişken seçim yöntemi kullanılmaktadır. Bu yöntemler, ileri seçim, geri seçim ve kademeli seçim yöntemleridir. Çalışmada model geliştirme aşamasında kademeli seçim yöntemi kullanılmıştır. Kademeli seçim yöntemi, ileri ve geri seçim yöntemleri arasında melez bir yaklaşım olarak tanımlanabilir. Bu süreçte, istatistiksel olarak en anlamlı değişken modele ilk önce girer. Bir sonraki adımda, izleyen değişkenler ileri seçim mantığıyla modele dahil edilir. Yeni bir değişken girdiğinde, modeldeki herhangi bir faktörün çıkarılıp çıkarılmayacağı geri seçim yöntemindeki gibi kontrol edilmektedir. Algoritma, değişkenlerin etkililiği bitinceye kadar çalışmaya devam etmektedir. Modele başka bir değişken eklenemez veya modele girilen değişken geriye dönük elemelerde çıkarılan değişken ise kademeli seçim süreci sona ermektedir.

Değişken seçim yöntemi belirlendikten sonra Regresyon düğümüyle parametre tahminleri yapılır. Parametre tahminleri yapıldıktan sonra, değişkenler arasındaki korelasyon incelenmeli ve belirlenen korelasyon eşliğini aşan değişkenler mevcutsa en anlamlısını modelde bırakıp diğerlerini elemek gerekir. Diğer taraftan, modele giren değişkenlerin işaretinin negatif olması gerekmektedir. Pozitif işaretli değişken mevcut ise bu değişkenlerin çıkarılması gerekmektedir. WOE_RM10 ve WOE_FINANSAL1 değişkenlerinin işaretinin pozitif olduğu görülmüştür. Bu iki değişken modelleme kapsamında çıkartılıp Regresyon düğümü tekrar çalıştırılmıştır.

Çalışmadaki modellemede kullanılan değişkenler Tablo 5.7’de görüldüğü gibidir.

Tablo 5.7. Lojistik regresyon modelleme girdi değişkenleri

Sıra	Değişken
1	WOE_GECIKME1
2	WOE_RM1
3	WOE_PROJE
4	WOE_RM8
5	WOE_LIMIT3
6	WOE_RM2
7	WOE_RM6
8	WOE_RM5
9	WOE_RM4
10	WOE_TEMINAT
11	WOE_GECIKME2
12	WOE_ORTAK
13	WOE_FINANSAL2
14	WOE_CEK2
15	WOE_CEK1
16	WOE_RM9

Bu değişkenlerden WOE_PROJE, WOE_RM2, WOE_TEMINAT, WOE_CEK2 ve WOE_RM9 değişkenleri istatistiki olarak anlamlı olmadıkları için elenmiştir. Bu safhada elenen değişkenler şöyledir:

Modelde kullanılan değişkenlerin açıklayıcılık gücünün önemli olduğuna önceki açıklamalarımızda değinilmişti. Çalışmada değişkenlerin giriş sırası aynı zamanda açıklayıcılık gücünü göstermektedir. Değişkenlerin giriş sırası Tablo 5.8’de gösterilmiştir.

Tablo 5.8. Lojistik regresyon analizinde deęişkenlerin modele giriş sırası

Adım	Modele Giren Deęişken	Serbestlik Derecesi	Score	
			Chi-Square	Pr > ChiSq
1	WOE_GECIKME1	1	1144,662	<,0001
2	WOE_RM4	1	364,7302	<,0001
3	WOE_RM1	1	172,3658	<,0001
4	WOE_LIMIT3	1	191,0488	<,0001
5	WOE_ORTAK	1	111,5487	<,0001
6	WOE_GECIKME2	1	92,2996	<,0001
7	WOE_RM5	1	89,1663	<,0001
8	WOE_RM8	1	60,0406	<,0001
9	WOE_RM6	1	31,3682	<,0001
10	WOE_FINANSAL2	1	14,5274	<,0001
11	WOE_CEK1	1	8,4715	0,0036

Tablo 5.8'de görüldüğü üzere açıklayıcı gücü en yüksek olan deęişkenler sırasıyla WOE_GECIKME1, WOE_RM4 ve WOE_RM1'dir.

Deęişkenlerin açıklayıcı gücü belirlendikten sonra modeli oluşturmak için parametre katsayılarının belirlenmesi gerekmektedir. Parametre katsayıları Tablo 5.9'da görüldüğü gibidir.

Tablo 5.9. Lojistik regresyon analizi parametre katsayıları

Parametreler	Serbestlik Derecesi	Kestirim Değeri	Standart Hata	Wald Chi-Square	Pr > ShiSq	Üstel Kestirim Değeri
Sabit Değer	1	-3,5024	0,0357	9640,96	<,0001	0,030
WOE_CEK1	1	-0,2042	0,0702	8,46	0,0036	0,815
WOE_FINANSAL2	1	-0,262	0,0702	13,91	0,0002	0,77
WOE_GECIKME1	1	-0,378	0,0427	78,45	<,0001	0,685
WOE_GECIKME2	1	-0,5516	0,0549	101	<,0001	0,576
WOE_LIMIT3	1	-0,4418	0,0536	68	<,0001	0,643
WOE_ORTAK	1	-0,5264	0,0579	82,65	<,0001	0,591
WOE_RM1	1	-0,4705	0,0427	121,19	<,0001	0,625
WOE_RM4	1	-0,4541	0,0608	55,85	<,0001	0,635
WOE_RM5	1	-0,3943	0,0475	68,78	<,0001	0,674
WOE_RM6	1	-0,3485	0,0658	28,02	<,0001	0,706
WOE_RM8	1	-0,2758	0,059	21,89	<,0001	0,759

Tablo 5.9’da görüldüğü gibi modelde kullanılacak katsayılar “Kestirim Değeri” kolonundadır.

Modelleme veri seti ile geliştirilmiş model örneğine ait skor kodları, nihai skoru hesaplamak için başlangıç veri setine uygulanmaktadır. Başka bir deyişle, skor, iyi ve kötü elemesi sonrasında müşterileri içeren model örneğine göre SAS Miner’da hesaplanmaktadır. Skor elde edildikten sonra, tüm müşterileri içeren kitleye kalibre edilir. Bu çalışma kapsamında kalibrasyon yapılmayacaktır.

Lojistik regresyon analizi uygulamasının detaylı SAS kodlarına Ek B’de yer verilmiştir.

5.3.5 Modelin performansı ve validasyonu

Yapılan çalışmalarda modelin performansını ölçmek için önceki açıklamalarda değinilen GINI katsayısının, ROC eğrisinin ve Hata Matrisi tablolarının değerlerinden yararlanarak bazı hesaplamaların yapılması gerekir. Bu hesaplamalara ilişkin değerler aşağıdaki Tablo 5.10'da görüldüğü gibidir.

Tablo 5.10. Lojistik regresyon modeline dair temel istatistikler

Uyum İyiliği İstatistikleri	İstatistik Adı	Öğrenme	Doğrulama
MISC	Misclassification Rate	0,029485257	0,029555789
KS	Kolmogorov-Smirnov Statistic	0,526259744	0,544860959
AUR	Area Under ROC	0,833601046	0,839553174
Gini	Gini Coefficient	0,667202092	0,679106348

Tablo 5.10'da görüldüğü gibi çalışmamızdaki GINI katsayısı % 66,72 olarak hesaplanmıştır. ROC değeri ise % 83,36 olarak elde edilmiştir. Hesaplanan bu değerler bize kurulan modelin yüksek tahmin gücüne sahip olduğunu göstermektedir.

Modelin performansının değerlendirilmesinde diğer bir analiz yöntemi olan kesim noktalarına göre hata matrisi tablolarının incelenmesi gerekir. Tablo 5.10'da görüldüğü üzere modelin yanlış sınıflama oranı % 2.94'tür.

Bu çalışmada temerrüt edenleri daha doğru sınıflandırabilmek için 5 farklı kesim noktasına göre analizler yapılmıştır. Kesim noktası olarak örneklemin temerrüt oranı 0,029, 0,05, 0,1, 0,20 değerleri ve genelde kullanılan bir değer olan 0,50 değeri alınmıştır. Analiz, regresyon düğümüne kesim noktası düğümü eklenerek yapılmıştır. Elde edilen sonuçlar aşağıda Tablo 5.11'de özetlenmiştir.

Tablo 5.11. Kesim noktalarına göre hata matris tabloları

Hata Matrisi 1			
Kesim Noktası: 0,029		Tahmin	
		Canlı	Temerrüt
Gerçek	Canlı	29.740	9.103
	Temerrüt	287	890

Hata Matrisi 2			
Kesim Noktası: 0,05		Tahmin	
		Canlı	Temerrüt
Gerçek	Canlı	33.750	5.093
	Temerrüt	448	729

Hata Matrisi 3			
Kesim Noktası: 0,1		Tahmin	
		Canlı	Temerrüt
Gerçek	Canlı	36.769	2.074
	Temerrüt	726	451

Hata Matrisi 4			
Kesim Noktası: 0,2		Tahmin	
		Canlı	Temerrüt
Gerçek	Canlı	38.246	597
	Temerrüt	969	208

Hata Matrisi 5			
Kesim Noktası: 0,5		Tahmin	
		Canlı	Temerrüt
Gerçek	Canlı	38.827	16
	Temerrüt	1.164	13

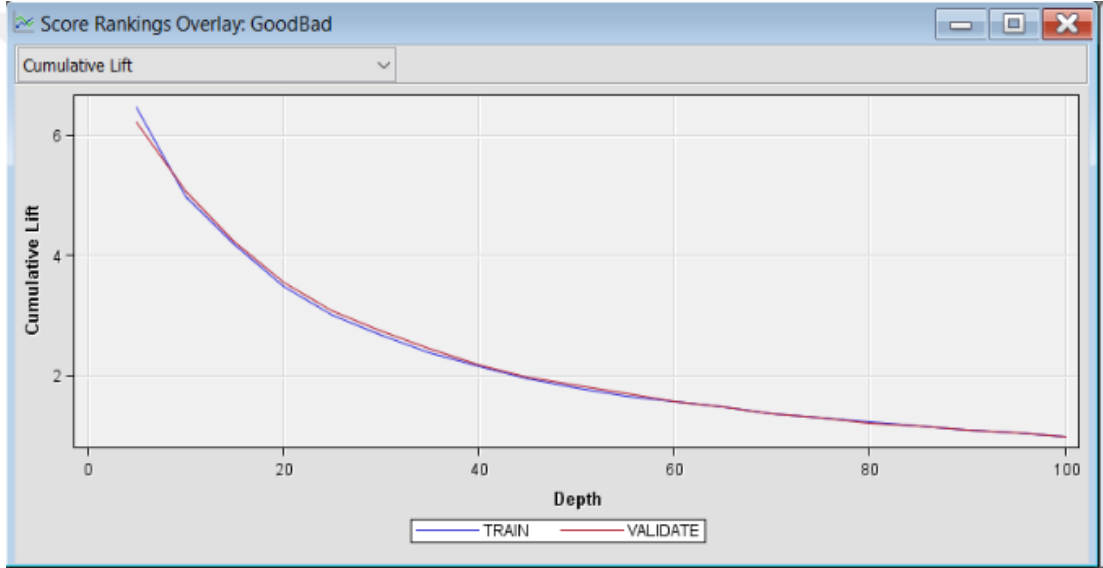
Tablolardan anlaşılacağı üzere, kesim noktası ne kadar küçük olursa temerrütler o kadar daha doğru tahmin edilebilmektedir. Diğer taraftan kesim noktası arttıkça temerrüt edecek müşterileri yanlış tahmin etme olasılığı da artmaktadır. Ancak, temerrütleri doğru tahmin etmenin bir maliyeti de olduğu bilinmelidir. Modina ve Pietrovito (2014: 1550)'nun belirttiği gibi, temerrüt edenleri doğru sınıflandırmak da çok önemlidir. Örneğin, temerrüt edecek bir müşteri yanlış sınıflandırıldığında, yani temerrüt etmeyecek olarak sınıflandırıldığında, banka o müşteriye verdiği kredi tutarı kadar zarar edebilir. Tersi durumda temerrüt etmeyecek bir müşteri temerrüt edecek olarak sınıflandırıldığında ise, banka o müşteriye kredi vermeyecek ve elde edeceği kârı kaybetmiş olacaktır. Ancak bu kâr tutarı genelde krediden çok düşük olmaktadır. Temerrüt edecek bir müşteriye yanlış tahmin etmek bankaya daha çok maliyet yüklemektedir.

Temerrüt edecek müşterileri yanlış tahmin etme olasılığı kesim noktasına bağlıdır. Kesim noktası arttıkça temerrüt eden müşterileri yanlış tahmin etme olasılığı da artmaktadır. Kesim noktasının belirlenmesinde ilgili finansal kuruluşun risk iştahı da belirleyici olmaktadır. Model ile her bir müşteri için temerrüt etme olasılığı hesaplanacağından, banka kendi belirlediği kesim noktasına göre müşterileri sınıflandırabilir ve içsel tanımlarına göre aksiyon alabilir.

Yukarıdaki açıklamalarda modelin performansının değerlendirilmeleri anlatılmıştır. Modelin performansı kadar validasyonu da incelenmelidir.

Validasyon (doğrulama), tahmin modeli geliştirmenin önemli bir aşamasıdır. Validasyon kurulan modelin sağlamlığını amaçlar ve modelin, model geliştirmede kullanılmayan örneklem üzerinde çalışıp çalışmadığını göstermektedir. Validasyon, SAS Miner Veri Bölme düğümü kullanılarak, istenilen yüzdeye bağlı bir rasgele örneklem seçilerek yapılmaktadır. Veri Bölme düğümü, tahmin modelini doğrulamak için başka bir rasgele örneklem oluşturmaya gerek olmaksızın yeni bir örneklemi otomatik olarak oluşturmaktadır.

Validasyon için, öğrenme ve doğrulama veri setleri incelenerek bu veri setlerinin sonuçlarının benzer olup olmadığı kontrol edilmelidir. Diğer bir ifade ile öğrenme ve doğrulama veri setlerinin sonuçları benzer olmalıdır. İlk önce GINI-ROC değerleri kontrol edilir. Öğrenme ve doğrulama veri setleri için birbirlerine yakın değerler olmalıdır. İkincisi, öğrenme ve doğrulama veri setleri ile kurulan modeller Kolmogorov-Smirnov istatistikleri açısından yakın değerler almalıdır. Tablo 5.10'da görüldüğü üzere, öğrenme ve doğrulama veri setleri bu iki kriterde yakın değerlere sahiptir. Üçüncüsü, bu iki veri seti ile elde edilen skor derecelendirmeleri birbirine yakın olmalıdır.



Şekil 5.1 Skor derecelendirme

Şekil 5.1'de görüldüğü üzere öğrenme ve doğrulama veri setleri ile elde edilen skor derecelendirmelerinin birbirine çok yakın olduğu görülmektedir.

5.4 Cox Regresyon Analizi Uygulaması

Cox regresyon analizi ile modellemeye başlamadan önce araştırmanın dizaynı ile ilgili temel hususlar gözden geçirilmeli ve araştırmanın başlangıç ve bitiş tarihleri ve kullanılacak zaman ölçeği kesin bir şekilde belirlenmeli ve başarı ve başarısızlık durumları net bir şekilde tanımlanmalıdır. Son yıllarda yapılan

çalıřmalarda Cox regresyon modeli sabit aıklayıcı deęiřkenlerin yanı sıra zamana baęlı aıklayıcı deęiřkenleri de ieren Cox regresyon modeline geniřlemiřtir.

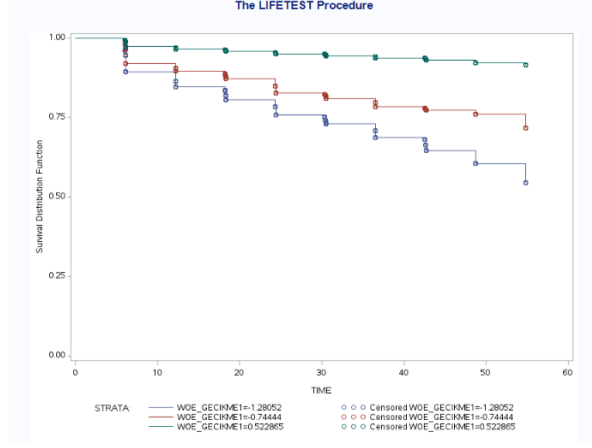
Lojistik regresyon analizi uygulamasında kullanılan veri seti dnřtrlerek Cox regresyon analizine uygun hale getirilmiř ve veri mřteri bazında yeniden dzenlenmiřtir. Mřterinin veri setinde ilk yer aldıęı dnem, o mřteri iin bařlangı tarihi olarak kabul edilmiř ve mřterinin veri setinde gzlemlendięi son dnem ise bitiř tarihi olarak alınmıřtır. Bu dnřm yapılırken Er (2010: 134)'in ikinci arařtırma dizaynına paralel bir yaklařım sergilenmiřtir. řyle ki temerrt durumlarında temerrtten nceki ilk dneme ait veriler ve temerrt yařanmadıęı durumlarda ise en gncel dneme ait veriler kullanılmıřtır. Bylece orijinal veri setinde 6'řar aylık olan ve zamanla deęiřen deęiřkenler, bir defa ieri alınmıřtır ve sadece son kesit tarihinde yer alan gzlemlerin 6 ay gzlemlendięi varsayılmıřtır.

Er (2010: 134)'in alıřmasında, bankaların 1989-2008 dnemine ait yıllık finansal oranları, banka bařarısızlıklarını aıklayıcı deęiřkenler olarak ele alınmıř ve saękalım analiziyle tahmin modelleri geliřtirilmiřtir. Arařtırma dizaynlarının birisinde zamana baęlı deęiřken olan yıllık finansal oranlar, dięerinde ise zamana baęlı deęiřken olan finansal oranların sadece belli bir yıla ait olanları kullanılarak modellenme yapılmıřtır. Hangi yıla ait finansal oranların kullanılacaęı hususunda bankaların kuruluř yılı baz alınmıřtır. Arařtırmanın bařlangı dneminden nce yani 1989'dan nce kurulan bankalar iin 1989 yılı referans alınmıřtır. zetle, "Veri setinin bařlangı tarihinde bankalara ait finansal oranlar kullanılarak da model tahminleri gerekleřtirilmiřtir. ... gerekleřtirilen model tahminlerinde ama bankanın ilk yıl performansının ve o yıl iinde bulunan makroekonomik kořulların bankanın yařam sresi zerindeki etkisini belirlemektir."

Cox regresyon modeli iin, tm deęiřkenlerin zamanla etkileřimleri modele deęiřken olarak dhil edilip bu zaman deęiřkenlerinin istatistiki olarak anlamlı

olup olmadıkları kontrol edilmiştir. Borucka (2014: 86), bu varsayımın önemine dair literatürde bir uzlaşının olmadığını, iki temel görüşün mevcut olduğunu belirtmektedir. Birinci görüş, bu varsayımın sağlanmamasını aşırı bir sorun olarak görmemekte ve değişkenin katsayısının araştırmaya konu zamansal noktaların 'ortalama etkisi' olarak yorumlanmasını önermektedir. İkinci görüş ise, varsayımın sağlanmamasını önemli bir sorun olarak görmekte ve modelde bazı modifikasyonların yapılmasını önermektedir. Bu modifikasyonlar iki farklı yöntemle yapılabilmektedir. Birinci yöntemde değişkenlerin zamanla etkileşimleri modele yeni birer değişken olarak eklenmektedir. İkinci yöntemde ise tabakalaştırma tercih edilmektedir. Tabakalaştırma yönteminde varsayımın sağlanmadığı kategorik değişkenin değerlerine göre, veri seti alt gruplara ayrılır ve her bir alt grup için ayrı tahmin modeli geliştirilir. Tabakalaştırmada kullanılan değişken, artık açıklayıcı bir değişken değildir. Ancak tabakalaştırmaya konu değişken sürekli bir değişken ise öncelikle değişkenin dönüştürülerek kategorik hale getirilmesi gerekmektedir.

Cox regresyon modelindeki değişkenlerin istatistiki olarak anlamlı olup olmadıkları varsayımı grafiksel yöntemlerle de değerlendirilebilir. Bu varsayım, grafiksel yöntemler ile değerlendirilirken iki kritere bakılmaktadır. Bu kriterlerden birincisi, ilgili değişkenin her bir kategorisi için oluşan trend eğrilerinin birbirine paralel olması, ikincisi ise, bu eğrilerin birbirleriyle kesişmemesidir. Grafiksel yöntem kullanılarak bakıldığında çoğu değişkenin bu varsayımı sağladığı görülmektedir. Fakat Şekil 5.2'de olduğu gibi, çalışmamızdaki bu varsayımı sağlamayan değişkenlerden birisi incelendiğinde trend eğrilerinin birbirine paralel olmadığı görülmektedir. Ancak grafiksel yöntemde yapılan yorumlar öznel olduğu için bu varsayımın testi için başka yöntemlerin de denenmesi gerekir.



Şekil 5.2 Grafiksel testler

Modeldeki değişkenlerin istatistiki olarak anlamlı olup olmadıkları varsayımını test etmek için ayrıca tüm değişkenlerin zamanla etkileşimleri ayrı birer değişken olarak ele alınmış ve zaman etkili bu değişkenlerin analizi yapılmıştır. Bu analiz Tablo 5.12’de görüldüğü gibidir.

Tablo 5.12. Zaman etkileşimli değişkenlerin analizi

Parametreler	Kestirim Değeri	Standart Hata	Wald Chi-Square	Pr > ShiSq	Üstel Kestirim Değeri
t*WOE_CEK1	-0,0147	0,028	0,28	0,5998	0,985
t*WOE_FINANSAL1	-0,0968	0,0481	4,05	0,0441	0,908
t*WOE_FINANSAL2	-0,0106	0,0424	0,06	0,8018	0,989
t*WOE_GECIKME1	-0,0539	0,0213	6,39	0,0115	0,948
t*WOE_LIMIT1	-0,0312	0,0361	0,75	0,3875	0,969
t*WOE_LIMIT2	-0,0256	0,0204	1,57	0,2106	0,975
t*WOE_LIMIT3	0,0169	0,0222	0,58	0,4474	1,017
t*WOE_MEVDUAT	0,0707	0,039	3,28	0,0701	1,073
t*WOE_ORTAK	0,1006	0,0329	9,33	0,0023	1,106
t*WOE_RM1	-0,0397	0,0227	3,05	0,0806	0,961
t*WOE_RM10	0,00592	0,0396	0,02	0,8812	1,006

Parametreler	Kestirim Deęeri	Standart Hata	Wald Chi-Square	Pr > ShiSq	Üstel Kestirim Deęeri
t*WOE_RM2	0,014	0,026	0,29	0,5904	1,014
t*WOE_RM4	-0,024	0,0408	0,34	0,5573	0,976
t*WOE_RM6	-0,0225	0,0291	0,6	0,4394	0,978
t*WOE_RM8	0,00885	0,0263	0,11	0,7369	1,009
t*WOE_RM9	0,0478	0,0422	1,29	0,2568	1,049
t*WOE_YAS2	0,00278	0,0519	0	0,9572	1,003

Tablo 5.12’de görüldüğü gibi zaman etkileşimli deęişkenlerin bazıları anlamlı bazıları ise anlamsızdır. Anlamlı çıkan bu deęişkenler daha önceki açıklamalarda deęinildiği gibi oransal tehlike (hazard) varsayımının ihlal edildiğini göstermektedir. Oransal tehlike varsayımını ihlal eden deęişkenler olduđu için Genişletilmiş Cox regresyon modeliyle modelleme yapmaya karar verilmiştir. Genişletilmiş Cox regresyon modeli, zamana bağımlı deęişkenlerin kullanılmasına izin vermektedir.

Genişletilmiş Cox regresyon modeli için hazırlanan veri setinde gözlemlerin dağılımı aşağıdaki Tablo 5.13’te görüldüğü gibidir.

Tablo 5.13. Genişletilmiş cox regresyon modelindeki gözlemlerin dağılımı

Bağımlı Deęişken Deęeri	Gözlem Adedi	Gözlem Yüzdesi
Canlı	22.853	93,3%
Temerrüt	1.643	6,7%
Toplam	24.496	100,0%

Genişletilmiş Cox regresyon modelindeki parametrelerini belirlemek amacıyla zamana bağımlı deęişkenleri içeren Cox regresyon analizi uygulanmıştır. Cox regresyon analiz uygulaması SAS Enterprise Guide 7.1 ve SAS Enterprise Miner

Client 4.1 ile yapılmıştır. Veri, Guide üzerinde modellemeye hazır hale getirilmiştir. Ardından Miner ile modelleme yapılmıştır.

Genişletilmiş Cox regresyon modelinin uygulama süreci veri setinin SAS Miner programına aktarılması ile başlamaktadır. 'StatExplore' düğümüyle betimleyici istatistiksel analizler yapıldıktan sonra, Veri Bölme düğümüyle veri seti ikiye ayrılmakta, ardından Etkileşimli Gruplama düğümüyle değişkenlere WoE dönüşümü uygulanmakta ve son olarak Sağkalım düğümüyle model geliştirilmektedir. Modelleme aşamasında, öncelikle tüm değişkenler incelenerek zaman etkileşimli değişkenler tespit edilir ve bunları içeren bir model geliştirilir. Bu aşamadan sonra kademeli seçim yöntemi kullanılarak Genişletilmiş Cox regresyon modeli kurulmaktadır.

5.4.1 Değişken analizleri

Cox regresyon analizi uygulaması için oluşturulan veri seti baz alınarak tek değişkenli analizler yapılmış ve anlamlı değişkenlerle modelleme süreci başlatılmıştır.

Etkileşimli gruplama öncesi değişkenler arasındaki maksimum pozitif korelasyon RM10 ile RM7 değişkenleri arasındadır ve ilgili korelasyon değeri 0,45654'tür. Etkileşimli gruplama öncesi değişkenler arasındaki maksimum negatif korelasyon GECIKME1 ile RM1 değişkenleri arasındadır ve ilgili korelasyon değeri -0,33686'dır.

Etkileşimli gruplama safhasında ise WoE dönüşümü yapılmış ve bazı değişkenler arasında aşırı korelasyon olduğu görülmüştür. WoE dönüşümü sonucunda PROJE, CEK2, TEMINAT ve RM5 değişkenlerinin WoE değerleri ile geri kalan diğer değişkenlerin WoE değerleri arasında yüksek korelasyon olduğu için bu değişkenler elenmiştir.

5.4.2 Etkileşimli gruplama

Etkileşimli gruplama yöntemi lojistik regresyon analizi bölümünde kapsamlı bir şekilde açıklanmıştır. Çalışmada etkileşimli gruplama düğümüyle değişkenlere monotonik bir yapı kazandırılmış ve değişken seçiminde aynı eşik değerler kullanılmıştır. Diğer bir ifade ile önceki açıklamalarda değinildiği gibi GINI'si 20'den küçük olan veya Bilgi Değeri 0,14'ten küçük olan değişkenler elenerek modellemeye alınmamıştır.

Modellemede kullanılan her bir değişken için GINI katsayısı ve Bilgi Değerleri hesaplanmıştır. Her bir değişken için hesaplanan GINI katsayısı ve Bilgi Değerleri Tablo 5.14'te görüldüğü gibidir.

Tablo 5.14. Genişletilmiş cox regresyon modelindeki değişkenlerin gini katsayıları ve bilgi değerleri

Sıra	Değişken	Hesaplanan Rol	GINI Katsayısı	Bilgi Değeri
1	LIMIT2	Kabul Edildi	43,22	0,70
2	LIMIT3	Kabul Edildi	43,15	0,65
3	RM8	Kabul Edildi	42,37	0,61
4	GECIKME1	Kabul Edildi	36,13	0,58
5	RM6	Kabul Edildi	34,72	0,41
6	RM1	Kabul Edildi	30,44	0,45
7	MEVDUAT	Kabul Edildi	27,80	0,26
8	RM10	Kabul Edildi	26,08	0,25
9	CEK1	Kabul Edildi	26,08	0,32
10	RM4	Kabul Edildi	25,10	0,21
11	ORTAK	Kabul Edildi	25,08	0,21
12	LIMIT1	Kabul Edildi	24,38	0,22
13	RM9	Kabul Edildi	23,31	0,18
14	RM2	Kabul Edildi	23,13	0,34
15	YAS2	Kabul Edildi	22,96	0,17

Sıra	Değişken	Hesaplanan Rol	GINI Katsayısı	Bilgi Değeri
16	FINANSAL2	Kabul Edildi	22,84	0,18
17	FINANSAL1	Kabul Edildi	21,37	0,15
18	RM3	Red Edildi	16,14	0,09
19	GECIKME2	Red Edildi	16,04	0,19
20	YAS1	Red Edildi	12,88	0,05
21	CEK3	Red Edildi	12,12	0,10
22	RM7	Red Edildi	8,76	0,03
23	GORUSME	Red Edildi	8,51	0,19
24	ISTIHBARAT	Red Edildi	1,05	0,00

Tablo 5.14'te görüldüğü gibi GINI ve Bilgi Değeri için belirlenen eşik değerlerden dolayı 7 tane değişken elenerek modelleme dışında bırakılmıştır. Elenen değişkenler, CEK3, GECIKME2, GORUSME, RM3, YAS1, RM7 ve ISTIHBARAT'tır.

5.4.3 Öğrenme ve doğrulama veri setleri

Öğrenme ve doğrulama setleri önceki açıklamalarda değinildiği gibi hesaplanmıştır. Yani Genişletilmiş Cox regresyon analizi için hazırlanan veri setinin % 70'i öğrenme ve % 30'u da doğrulama için ayrılmıştır. Bu işlem, Veri Bölme düğümüyle yapılmıştır. Öğrenme ve doğrulama veri setlerinin sonuçları Tablo 5.15'te görüldüğü gibidir.

Tablo 5.15. Genişletilmiş cox regresyon öğrenme ve doğrulama veri setleri

Öğrenme Veri Seti		
Bağımlı Değişken Değeri	Gözlem Adedi	Gözlem Yüzdesi
Canlı	15.996	93,3%
Temerrüt	1.149	6,7%
Toplam	17.145	100,0%

Doğrulama Veri Seti		
Bağımlı Değişken Değeri	Gözlem Adedi	Gözlem Yüzdesi
Canlı	6.857	93,3%
Temerrüt	494	6,7%
Toplam	7.351	100,0%

Tablo 5.15'te görüldüğü gibi öğrenme ve doğrulama veri setlerinin her ikisinin de temerrüt oranları aynıdır.

5.4.4 Değişken seçimi ve regresyon

Model geliştirme aşamasında lojistik regresyonda olduğu gibi kademeli seçim yöntemi tercih edilmiştir.

Kademeli seçim yöntemi sonucunda elde edilen değişkenler aşağıda görüldüğü gibidir. WOE_CEK1, WOE_FINANSAL2, WOE_GECIKME1, WOE_GECIKME2, WOE_GORUSME, WOE_LIMIT1, WOE_LIMIT2, WOE_LIMIT3, WOE_MEVDUAT, WOE_ORTAK, WOE_RM1, WOE_RM2, WOE_RM6, WOE_RM8 ve WOE_YAS2 değişkenleri modele girmiştir. Diğer taraftan, zaman değişkeni ile etkileşimi anlamlı olan $t \cdot WOE_GECIKME1$ ve $t \cdot WOE_ORTAK$ değişkenleri de modelde yer almıştır. Bazı değişkenler ise, istatistiki olarak anlamlı olmadıkları için elenmiştir. Bu değişkenler, FINANSAL1, WOE_RM4, WOE_RM9 ve WOE_RM10'dur.

Çalışmamızdaki modele dâhil olan değişkenlerin giriş sırası aynı zamanda değişkenlerin açıklayıcılık gücünü de gösterdiği önceki açıklamalarda belirtilmişti. Değişkenlerin modele giriş sırası Tablo 5.16'da gösterildiği gibidir.

Tablo 5.16. Genişletilmiş cox regresyondaki değişkenlerin modele giriş sırası

Adım	Modele Giren Değişken	Serbestlik Derecesi	Score Chi-Square	Pr > ChiSq
1	WOE_LIMIT2	1	1.259,30	<.0001
2	WOE_GECIKME1	1	661,48	<.0001
3	WOE_YAS2	1	431,21	<.0001
4	WOE_LIMIT3	1	298,02	<.0001
5	WOE_MEVDUAT	1	229,90	<.0001
6	WOE_RM2	1	139,95	<.0001
7	WOE_RM1	1	94,23	<.0001
8	WOE_ORTAK	1	73,36	<.0001
9	WOE_RM8	1	58,47	<.0001
10	WOE_LIMIT1	1	21,91	<.0001
11	WOE_RM6	1	15,24	<.0001
12	_t_*WOE_GECIKME1	1	13,91	0.0002
13	WOE_CEK1	1	7,73	0.0054
14	_t_*WOE_ORTAK	1	7,76	0.0053
15	WOE_FINANSAL2	1	4,62	0.0316

Değişkenlerin sırası ve açıklayıcılık gücü açıklandıktan sonra regresyon modelini oluşturmamız için parametre katsayılarının hesaplanması gerekmektedir. Genişletilmiş Cox regresyon parametrelerinin katsayıları Tablo 5.17'de görüldüğü gibidir.

Tablo 5.17. Genişletilmiş cox regresyon analizi parametre katsayıları

Parametreler	Serbestlik Derecesi	Kestirim Değeri	Standart Hata	Wald Chi-Square	Pr > ShiSq	Üstel Kestirim Değeri
WOE_CEK1	1	-0,15	0,0573	6,85	0,0088	0,861
WOE_FINANSAL2	1	-0,1792	0,0835	4,61	0,0318	0,836
WOE_GECIKME1	1	-0,1896	0,0654	8,41	0,0037	0,827
WOE_LIMIT1	1	-0,3485	0,078	19,94	<,0001	0,706
WOE_LIMIT2	1	-0,8581	0,0443	374,89	<,0001	0,424
WOE_LIMIT3	1	-0,5263	0,0467	127,12	<,0001	0,591
WOE_MEVDUAT	1	-0,8878	0,07	160,84	<,0001	0,412
WOE_ORTAK	1	-0,7203	0,1068	45,52	<,0001	0,487
WOE_RM1	1	-0,3695	0,0478	59,81	<,0001	0,691
WOE_RM2	1	-0,3443	0,0495	48,4	<,0001	0,709
WOE_RM6	1	-0,2351	0,0613	14,71	0,0001	0,79
WOE_RM8	1	-0,2482	0,0528	22,12	<,0001	0,78
WOE_YAS2	1	-17,526	0,1037	285,64	<,0001	0,173
t*WOE_GECIKME1	1	-0,0716	0,0185	15,01	0,0001	0,931
t*WOE_ORTAK	1	0,0882	0,0316	7,79	0,0053	1,092

Tablo 5.17’de görüldüğü gibi kestirim değerleri bize regresyon modelinde kullanılacak parametrelerin katsayılarını vermektedir.

Genişletilmiş Cox regresyon analizi uygulamasının detaylı SAS çıktılarına Ek C’de yer verilmiştir.

5.4.5 Modelin performansı ve validasyonu

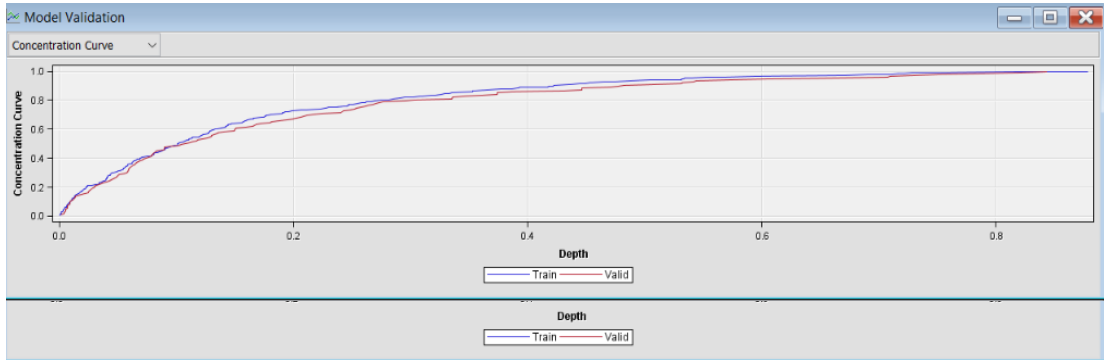
Modelin performansını ve validasyonunu incelemek için önceki açıklamalarımızda değinildiği gibi temel istatistiklere bakılması gerekir. GINI ve Kolmogorov-Smirnov gibi istatistiklerinin hesaplanması gerekmektedir. Hesaplanan bu değerlere Tablo 5.18’de görüldüğü gibidir.

Tablo 5.18. Geniřletilmiř cox regresyon modeline dair temel istatistikler

Veri Seti	Benefit	Average Hazard Function	Depth	Lift	Kolmogorov-Smirnov Statistic	Gini Concentration Ratio
Öğrenme	0,52	0,0195	0,209	3,5	0,54	0,6674
Doğrulama	0,49	0,0190	0,225	3,2	0,51	0,6084

Modelin performansı açısından deęerlendirme yapılırken doğrulama ve öğrenme veri setlerinin GINI deęerlerine bakılır. Tablo 5.18’de görüldüğü gibi öğrenme ve doğrulama veri setlerinin GINI deęerleri birbirlerine oldukça yakın yüksek deęerlere sahiptir. Bu sonuç bize modelin performansının iyi olduğunu göstermektedir.

Diđer taraftan, validasyonun doğru yapıldığından emin olmak için 3 ana faktör deęerlendirmeye alınmaktadır. Birinci deęerlendirmede GINI deęerleri kontrol edilmektedir. Öğrenme ve doğrulama veri setlerinin GINI deęerleri birbirine yakın olmalıdır. İkinci deęerlendirmede, öğrenme ve validasyon veri setleri ile kurulan modeller, Kolmogorov-Smirnov istatistikleri açısından yakın deęerler almalıdır. Tablo 5.18 incelendiğinde bu iki kriterin sağlandığı görülmektedir. Üçüncü deęerlendirmede bu iki veri seti ile elde edilen validasyon sonuçları birbirine yakın olmalıdır. Öğrenme ve doğrulama veri setleri ile elde edilen validasyon sonuçları Şekil 5.3’te görüldüğü gibidir.





Şekil 5.3. Validasyon sonuçları

Şekil 5.3'te görüldüğü gibi iki veri seti ile elde edilen validasyon sonuçları birbirine çok yakındır.

6. SONUÇ VE ÖNERİLER

Geliştirilen modeller istatistiki olarak ve iş mantığı açısından anlamlıdır. Çalışmaya veri sağlayan banka, çalışma sonucunda geliştirilen regresyon modellerini kullanarak ilgili portföy için temerrüt olasılığı tahmini yapabilecektir. Geliştirilen modeller, beklenen kredi zararı hesaplamalarında, sermaye yeterliliği rasyosu hesaplamalarında ve kredi risk yönetiminde kullanılabilir. Risk yönetimi açısından, modelin çalıştırıldığı raporlama tarihinden sonraki bir yıllık süreçte, aktif kalitesine ilişkin çeşitli öngörüler yapılabilecektir. Modellerde geliştirilen temerrüt olasılıkları, içsel derecelendirme notuna çevrilerek bankanın kredi kalitesi net bir şekilde ortaya konabilir. Kredilerin yüksek içsel derecelendirme notlarında yoğunlaşması, bankanın kredi kalitesinin yüksek olduğunun göstergesidir. İçsel derecelendirme notları bankanın kredi kalitesini ölçmede kullanılabilir. Böylece bankalar kredi kalitesi düşen kredilerde daha erken aksiyon alabilirler, krediyi teminatlandırma olasılıklarını arttırabilir veya ilave kredi riskine girmeyebilirler.

Öğrenme veri seti ile geliştirilen lojistik regresyon modelinin GINI katsayısı, 0,6672'dir. Öğrenme veri seti ile geliştirilen Cox regresyon modelinin GINI katsayısı, 0,6674'tür. Lojistik regresyon modeli ile Genişletilmiş Cox Regresyon modeli, GINI değeri açısından karşılaştırıldığında birbirlerine çok yakın değerler elde edilmiştir. Çok az bir farkla Genişletilmiş Cox regresyon modeli daha açıklayıcı bulunmuştur. Kurulan modellerin tahmin güçleri birbirine yakındır. Bu durumda Genişletilmiş Cox regresyonun alternatif faydaları dikkate alınarak daha uygun bir teknik olduğu iddia edilebilir.

Çalışmada geliştirilen modeller, modellerin kurulduğu zaman dilimindeki her bir kesit tarihinde gözlemlenen temerrüt oranlarının ortalamasını yansıtmaktadır. Diğer bir ifade ile üretilen temerrüt olasılıkları, döngüsel temerrüt olasılıklarıdır. İDD yaklaşımı ile sermaye yeterliliği hesaplamasında döngüsel temerrüt olasılıkları kullanılmaktadır. TFRS 9, beklenen kredi zararı hesaplamalarına

güncel ekonomik durumun ve ileriye dönük tahminlerin yansıtılmasını gerektirmektedir. İleriye dönük tahminler, makroekonomik modelleme ile yansıtılmaktadır. Güncel ekonomik durumun müşteriler için hesaplanan temerrüt olasılıklarına yansıtılabilmesi için geliştirilen tahmin modellerinde bazı modifikasyonların yapılması gerekmektedir. Döngüsel temerrüt olasılıkları, bu modifikasyonlarla anlık temerrüt olasılıklarına dönüştürülmektedirler. Toffano vd. (2017) bu dönüşüm için farklı yöntemlerin olduğunu belirttikten sonra kendi yaklaşımlarını anlatmışlardır. Uygulamada genellikle kullanılan çözüm, geliştirilen modelin güncel bir döneme kalibre edilmesidir. Bu kalibrasyon işlemiyle anlık temerrüt olasılıkları elde edilmektedir.

Modeller geliştirilirken 2011 ile 2015 arasındaki veriler kullanılmıştır. Çalışmanın güncel dönemi yani ekonominin içinde bulunduğu mevcut konjonktürü daha kesin yansıtabilmesi için uygulamanın, güncel veriler eklenerek yapılması faydalı olacaktır. Modelleme, doğası gereği sürekli güncelleme gerektiren bir süreçtir. Geliştirilen modeli, yıllarca kullanılacak sabit bir model olarak görmek yerine, sürekli gözden geçirilmesi ve modifiye edilmesi gereken bir yapı olarak görmek gerekmektedir.

KAYNAKLAR

- Agresti, A., 2013. Categorical Data Analysis (Third Edition). John Wiley & Sons, Inc., 714, New Jersey.
- Andreeva, G., 2006. European Generic Scoring Models Using Survival Analysis. The Journal of the Operational Research Society, 57(10), 1180-1187.
- Alpar, R., 2003. Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş 1 (Değiştirilmiş ve Genişletilmiş 2. Baskı). Nobel Yayın Dağıtım, 411, Ankara.
- Alpar, R., 2017. Uygulamalı Çok Değişkenli İstatistiksel Yöntemler (5. Baskı). Detay Yayıncılık, 840, Ankara.
- Anavatan, A., 2011. Zaman Bağımlı Cox Regresyonu: Genç Firmaların Hayatta Kalma Başarısı Üzerine Bir Uygulama. Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, 154, Antalya.
- Aytürk, Y., 2016. UFRS 9 Finansal Araçlar Kapsamında Yeni Değer Düşüklüğü Modeli ve Bankacılık Sektörüne Etkileri. Mali Çözüm, 137, 133-143.
- BDDK, 2018. Kredilerin Sınıflandırılması Ve Bunlar İçin Ayrılacak Karşılıklara İlişkin Usul Ve Esaslar Hakkında Yönetmelik. Erişim Tarihi: 25.12.2018. <http://www.bddk.org.tr/Mevzuat-Kategori/Bankacilik-Kanuna-Iliskin-Duzenlemeler/2>.
- Bellotti, T., Crook, J., 2009. Credit Scoring with Macroeconomic Variables Using Survival Analysis. The Journal of the Operational Research Society, 60(12), 1699-1707.
- Borucka, J., 2014. Extensions of Cox Model for Non-Proportional Hazards Purpose. Econometrics, 3(45), 85-101.
- Börüban, C., 2009. Firmaların Mali Başarısızlıklarının Öngörülmesinde Diskriminant Analizi Ve Lojistik Regresyon Analizi Yöntemlerinin Karşılaştırılması. Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, 92, İstanbul.
- Bottazzi, G., Grazzi, M., Secchi, A., Tamagni, F., 2011. Financial and Economic Determinants of Firm Default. Journal of Evolutionary Economics, 373-476.
- Bülbül, S.E., 1999. İstatistiksel Başarısızlık Zamanı Modelleri Ve Finansal Analizlerde Uygulaması. Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Doktora Tezi, 216, İstanbul.

- Çokluk, Ö., 2010. Lojistik Regresyon Analizi: Kavram ve Uygulama. Kuram ve Uygulamada Eğitim Bilimleri, 10(3), 1357-1407.
- Çokluk, Ö., Şekercioğlu, G., Büyüköztürk, Ş., 2014. Sosyal Bilimler İçin Çok Değişkenli İstatistik SPSS Ve LISREL Uygulamaları (3. Baskı). Pegem Akademi, 414, Ankara.
- De George, E.T., Li, X., Shivakumar, L., September 2016. A Review of the IFRS Adoption Literature. Review of Accounting Studies, 21(3), 898-1004.
- Dominiak, P., Mazurkiewicz, M., 2011. Analysis of the Risk of Company's Bankruptcy in Polish Food and Beverage Production Sector Using the Cox Regression. Operations Research and Decisions, (1), 19-31.
- Er, E., 2010. Banka Başarısızlıklarının Sağkalım Analiziyle Açıklanması: Türkiye Örneği (1989–2008). Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, 168, Ankara.
- George, B., Seals, S., Aban, I., 2014. Survival Analysis and Regression Models. Journal of Nuclear Cardiology, 21(4), 686-694.
- Im, J-K., Apley, DW., Qi, C., Shan, X., March 2012. A Time-dependent Proportional Hazards Survival Model for Credit Risk Analysis. The Journal of the Operational Research Society, 63(3), 306-321.
- Işık, A., 2007. COX Regresyon Modeli İle Finansal Başarısızlığın Belirlenmesi Üzerine Bir Uygulama. Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, 203, İstanbul.
- Harrell, F.E., 2001. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis. Springer, 568, New York.
- Hosmer, D.W., Lemeshow, S., Sturdivant, R.X., 2013. Applied Logistic Regression (Third Edition). John Wiley & Sons, Inc., 500, New Jersey.
- Kamu Gözetim Kurumu, 2018. TFRS 9 Finansal Araçlar (2017 Sürümü). Erişim Tarihi: 17.01.2018.
<http://www.kgk.gov.tr/DynamicContentDetail/7890/TMS/TFRS-2018-Seti>.
- Karagöz, Y., 2016. SPSS 23 ve AMOS 23 Uygulamalı İstatistiksel Analizler. Nobel Akademik Yayıncılık, 1196, Ankara.
- Kavcıoğlu, Ş., 2011. Ticari Bankacılıkta Kredi Riskinin ve Kredi Riski Ölçüm Modellerinin Değerlendirilmesi. Finansal Araştırmalar ve Çalışmalar Dergisi, 3(5), 11-19.

- Kleinbaum, D.G., Klein, M., 2002. Logistic Regression A Self-learning Text. Springer, 513, New York.
- Kleinbaum, D.G., Klein, M., 2005. Survival Analysis A Self-Learning Text (Second Edition). Springer, 590, New York.
- Kocabaş, E., 2014. Lojistik Regresyon Ve Bankacılık Verileri Üzerine Bir Uygulama. Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 61, İstanbul.
- Laitinen, T., Kankaanpaa, M., 1999. Comparative Analysis of Failure Prediction Methods: the Finnish Case. European Accounting Review, 8(1), 67-92.
- Louzada, F., Cancho, V.G., Oliveira, M.R., Yiqi, B., 2014. Modeling Time to Default on a Personal Loan Portfolio in Presence of Disproportionate Hazard Rates. Journal of Statistics Applications & Probability, 3(3), 295-305.
- Marimo, M., Chimedza, C., 2017. Survival Analysis of Bank Loans in the Presence of Long-Term Survivors. South African Statistical Journal, (51), 199-216.
- McDonald, J.H., 2014. Handbook of Biological Statistics. Erişim Tarihi: 23.12.2018. <http://www.biostathandbook.com/independence.html> (Kitabın basılı nüshasının bilgileri şöyledir: McDonald, J.H., 2014. Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, 299, Baltimore.)
- Modina, M., Pietrovito, F., 2014. A Default Prediction Model for Italian SMEs: The Relevance of the Capital Structure. Applied Financial Economics, 24(23), 1537-1554.
- Özdemir, H., 2010. Lojistik Regresyon İle Kredi Skorlama Ve Bir Uygulama. Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 120, İstanbul.
- Peng, C.J., Lee, K.L., Ingersoll, G.M., 2002. An Introduction to Logistic Regression Analysis and Reporting, The Journal of Educational Research, 96(1), 3-14.
- Petrucci, C.J., 2009. A Primer for Social Worker Researchers on How to Conduct a Multinomial Logistic Regression. Journal of Social Service Research, 35(2), 193-205.
- Ptak-Chmielewska, A., 2016. Statistical Models For Corporate Credit Risk Assessment – Rating Models. Folia Oeconomica, 3(322), 87-111.
- Saltoğlu, M., Mart 2016. UFRS 9 ve Beklenen Kredi Zararları Modeli. Muhasebe Bilim Dünyası Dergisi, 18(1), 51-67.

- Sıddıqı, N., 2006. Credit Risk Scorecards Developing and Implementing Intelligent Credit Scoring. John Wiley & Sons, Inc., 196, New Jersey.
- Sıddıqı, N., 2017. Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards (2nd Edition). John Wiley & Sons, Inc., 464, New Jersey.
- Stepanova, M., Thomas, L.C., 2001. PHAB Scores: Proportional Hazards Analysis Behavioural Scores. The Journal of the Operational Research Society, Special Issue: Credit Scoring and Data Mining, 52(9), 1007-1016.
- Tabachnick, B.G., Fidel, L.S., 2015. Çok Değişkenli İstatistiklerin Kullanımı (Altıncı Basımdan Çeviri). Çev. Ed. Baloğlu, M., Nobel Akademik Yayıncılık, 1083, Ankara.
- Toffano, C., Nisi, F., Maurri, L., 2017. IFRS 9, Stress Testing, ICAAP: a comprehensive framework for PD calculation. Erişim Tarihi: 01/05/2019. <https://www.prometeia.it/en/media/ifrs9-stress-testing-icaap-comprehensive-framework-pd-calculation-prometeia-paper>.
- TBB, 2018. Bankalarca Kredilerin ve Diğer Alacakların Niteliklerinin Belirlenmesi ve Bunlar için Ayrılacak Karşılıklara İlişkin Usul ve Esaslar Hakkında Yönetmelik. Erişim Tarihi: 25/12/2018. <https://www.tbb.org.tr/tr/bankacilik/duzenlemeler/bankacilik-kanunu%E2%80%99na-iliskin-alt-duzenlemeler/75>
- Tezergil, S., 2009. Çok Aşamalı Lojistik Modellerin Finans Sektörüne Bir Uygulaması. Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Doktora Tezi, 174, İstanbul.
- Topcu, A., 2013. Cox Modeli İle Türkiye İmalat Sektöründeki Finansal Başarısızlığın Analizi. Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 136, Ankara.
- Turanlı, M., Güriş, S., Cengiz, D., Özden, Ü.H., Başar, Ö.D., Kalkan, S.B., 2017. İstatistik El Kitabı. Der Yayınları, 188, İstanbul.
- Treewichayapong, S., Chunchachinda, P., Padungsaksawasdi, C., 2011. Bankruptcy Prediction of Real Estate Firms in Thailand. The International Journal of Finance, 23(1), 6672-6691.
- UFRS Vakfı, 2018. About us. Erişim Tarihi: 01.06.2018. <https://www.ifrs.org/about-us/who-we-are/>

EKLER

Ek A. Değişkenler Listesi

Değişken Adı	Tanım
CEK1	Son 1 Yıl Teminata Alınan ve Karşılıksız Çıkan Çek Adedinin Son Bir Yıl Teminata Alınan Tüm Çeklere Oranı
CEK2	Son 1 Yıl Teminat Alınan Çeklerin Tutar Ortalaması
CEK3	Son 5 Yılda Kredi Risk Grubu Bazında Karşılıksız Çıkan Çek Adedi
FINANSAL1	VÖK / Toplam Banka Borçları
FINANSAL2	(Özkaynaklar + Ortaklara Borçlar) / (Toplam Borçlar - Ortaklara Borçlar - Hakedişler-Alınan Avanslar)
GECIKME1	Son 6 Ay İçindeki Maksimum Gecikme Gün Sayısı
GECIKME2	Kredi Risk Grubu Bazında Son Bir Aydaki Gecikme Gün Sayısı
GORUSME	Son 6 Ay İçinde Tahsilat Amaçlı Görüşme Adedi
ISTIHBARAT	Kredi Risk Grubu Bazında Gelen Resmi Yazı Adedi
LIMIT1	Önceki Son 12 Aydaki Ortalama Nakdi Limiti
LIMIT2	Müşteriye Son Limit Açılma Tarihi Üzerinden Geçen Süre
LIMIT3	Nakdi Limit Kullanım Oranı
MEVDUAT	Cari Hesaplarında Son 3 Ay Ortalama Bakiye
ORTAK	Son 3 Ayda Ana Ortağın Kredi Başvuru Sayısı
PROJE	Son 24 Ayda Yapılan Nakdi Proje Sayısı
RM1	Karpayı Taahhuklarda Değişim T0_T3
RM2	Faktoring Riskinde Değişim T0_T12
RM3	Toplam Risk Değişimi T0_T6
RM4	Çalışılan Banka Sayısı
RM5	Faktoring Riski - Toplam Risk Oranı T0
RM6	Toplam Limit Kullanımı T6
RM7	GN Limit Değişimi T0_T6
RM8	KV Nakdi Risk - Toplam Limit Oranı T0
RM9	UV Nakdi Risk Değişimi T0_T3
RM10	Toplam GN Risk Yüzdesi T0
TEMINAT	Mevcut Farklı Teminat Türü Sayısı
YAS1	Firma Yaşı
YAS2	Müşteriye Hesap Açılma Tarihi Üzerinden Geçen Süre

Ek B. Lojistik Regresyon Analizi SAS Çıktıları

```
Optimized SAS Code
1  *-----*
2  * EM SCORE CODE;
3  *-----*
4  *-----*
5  * TOOL: Input Data Source;
6  * TYPE: SAMPLE;
7  * NODE: Ids;
8  *-----*
9  *-----*
10 * TOOL: Statistics Exploration;
11 * TYPE: EXPLORE;
12 * NODE: Stat;
13 *-----*
14 *-----*
15 * TOOL: Partition Class;
16 * TYPE: SAMPLE;
17 * NODE: Part;
18 *-----*
19 *-----*
20 * TOOL: Extension Class;
21 * TYPE: CREDSCORE;
22 * NODE: IGN;
23 *-----*
24 length _UFormat $200;
25 drop _UFormat;
26 _UFormat='';
27
28 *-----*
29 * Variable: CEK1;
30 *-----*
31 LABEL WOE_CEK1 =
32 "Weight of Evidence: BdCqCtlAamtLtYrRtOvAllCqCtlLtYr";
33
34 if MISSING(CEK1) then do;
35 WOE_CEK1 = 0.3131322477;
36 end;
37 else if NOT MISSING(CEK1) then do;
38 if CEK1 < 7777 then do;
39 WOE_CEK1 = -0.86596691;
40 end;
41 else
42 if 7777 <= CEK1 AND CEK1 < 9999 then do;
43 WOE_CEK1 = -0.303035032;
44 end;
45 else
46 if 9999 <= CEK1 then do;
47 WOE_CEK1 = 0.3131322477;
48 end;
49 end;
50
51 *-----*
52 * Variable: FINANSAL2;
53 *-----*
54 LABEL WOE_FINANSAL2 = "Weight of Evidence: FINANSAL2";
55
56 if MISSING(FINANSAL2) then do;
57 WOE_FINANSAL2 = 0.988324968;
58 end;
59 else if NOT MISSING(FINANSAL2) then do;
60 if FINANSAL2 < 0.38 then do;
61 WOE_FINANSAL2 = -0.493035127;
62 end;
63 else
64 if 0.38 <= FINANSAL2 AND FINANSAL2 < 0.93 then do;
65 WOE_FINANSAL2 = 0.0618455579;
66 end;
67 else
68 if 0.93 <= FINANSAL2 AND FINANSAL2 < 1.79 then do;
69 WOE_FINANSAL2 = 0.5038894243;
70 end;
```

Optimized SAS Code

```

71     else
72     if 1.79 <= FINANSAL2 then do;
73     WOE_FINANSAL2 = 0.7824531689;
74     end;
75     end;
76
77     *-----*;
78     * Variable: GECIKME1;
79     *-----*;
80     LABEL WOE_GECIKME1 = "Weight of Evidence: GECIKME1";
81
82     if MISSING(GECIKME1) then do;
83     WOE_GECIKME1 = 0.6841006398;
84     end;
85     else if NOT MISSING(GECIKME1) then do;
86     if GECIKME1 < 1 then do;
87     WOE_GECIKME1 = 0.6841006398;
88     end;
89     else
90     if 1 <= GECIKME1 AND GECIKME1 < 3 then do;
91     WOE_GECIKME1 = -0.250035004;
92     end;
93     else
94     if 3 <= GECIKME1 AND GECIKME1 < 5 then do;
95     WOE_GECIKME1 = -0.576711476;
96     end;
97     else
98     if 5 <= GECIKME1 AND GECIKME1 < 16 then do;
99     WOE_GECIKME1 = -1.025103933;
100    end;
101    else
102    if 16 <= GECIKME1 then do;
103    WOE_GECIKME1 = -1.597168919;
104    end;
105    end;
106
107    *-----*;
108    * Variable: GECIKME2;
109    *-----*;
110    LABEL WOE_GECIKME2 = "Weight of Evidence: GECIKME2";
111
112    if MISSING(GECIKME2) then do;
113    WOE_GECIKME2 = 0.243235718;
114    end;
115    else if NOT MISSING(GECIKME2) then do;
116    if GECIKME2 < 1 then do;
117    WOE_GECIKME2 = 0.243235718;
118    end;
119    else
120    if 1 <= GECIKME2 AND GECIKME2 < 5 then do;
121    WOE_GECIKME2 = -0.49473048;
122    end;
123    else
124    if 5 <= GECIKME2 AND GECIKME2 < 21 then do;
125    WOE_GECIKME2 = -1.230229376;
126    end;
127    else
128    if 21 <= GECIKME2 then do;
129    WOE_GECIKME2 = -1.766821228;
130    end;
131    end;
132
133    *-----*;
134    * Variable: LIMIT3;
135    *-----*;
136    LABEL WOE_LIMIT3 =
137    "Weight of Evidence: CashLimitUtil";
138
139    if MISSING(LIMIT3) then do;
140    WOE_LIMIT3 = 0.5835334874;

```

Optimized SAS Code

```

141 end;
142 else if NOT MISSING(LIMIT3) then do;
143 if LIMIT3 < 0.02 then do;
144 WOE_LIMIT3 = 1.0839884341;
145 end;
146 else
147 if 0.02 <= LIMIT3 AND LIMIT3 < 0.46 then do;
148 WOE_LIMIT3 = 0.0132977457;
149 end;
150 else
151 if 0.46 <= LIMIT3 AND LIMIT3 < 0.75 then do;
152 WOE_LIMIT3 = -0.534112756;
153 end;
154 else
155 if 0.75 <= LIMIT3 then do;
156 WOE_LIMIT3 = -0.927512654;
157 end;
158 end;
159
160 *-----*;
161 * Variable: ORTAK;
162 *-----*;
163 LABEL WOE_ORTAK =
164 "Weight of Evidence: AvailblApplicantCountThreeMonth1";
165
166 if MISSING(ORTAK) then do;
167 WOE_ORTAK = 0.7140299763;
168 end;
169 else if NOT MISSING(ORTAK) then do;
170 if ORTAK < 1 then do;
171 WOE_ORTAK = 0.4187733882;
172 end;
173 else
174 if 1 <= ORTAK AND ORTAK < 2 then do;
175 WOE_ORTAK = 0.0204347671;
176 end;
177 else
178 if 2 <= ORTAK AND ORTAK < 3 then do;
179 WOE_ORTAK = -0.411463098;
180 end;
181 else
182 if 3 <= ORTAK AND ORTAK < 4 then do;
183 WOE_ORTAK = -0.734786814;
184 end;
185 else
186 if 4 <= ORTAK then do;
187 WOE_ORTAK = -1.039075616;
188 end;
189 end;
190
191 *-----*;
192 * Variable: RMI;
193 *-----*;
194 LABEL WOE_RMI =
195 "Weight of Evidence: ChangingIntAccuralTOT3";
196
197 if MISSING(RMI) then do;
198 WOE_RMI = -0.10749168;
199 end;
200 else if NOT MISSING(RMI) then do;
201 if RMI < 2222 then do;
202 WOE_RMI = -1.281651371;
203 end;
204 else
205 if 2222 <= RMI AND RMI < 6666 then do;
206 WOE_RMI = -0.872363161;
207 end;
208 else
209 if 6666 <= RMI then do;
210 WOE_RMI = 0.4844342639;

```


Optimized SAS Code

```

211 end;
212 end;
213
214 *-----*;
215 * Variable: RM4;
216 *-----*;
217 LABEL WOE_RM4 =
218 "Weight of Evidence: WorkingBankCount";
219
220 if MISSING(RM4) then do;
221 WOE_RM4 = -0.194568321;
222 end;
223 else if NOT MISSING(RM4) then do;
224 if RM4 < 8 then do;
225 WOE_RM4 = 0.5795465946;
226 end;
227 else
228 if 8 <= RM4 AND RM4 < 13 then do;
229 WOE_RM4 = 0.0421358352;
230 end;
231 else
232 if 13 <= RM4 AND RM4 < 18 then do;
233 WOE_RM4 = -0.515940416;
234 end;
235 else
236 if 18 <= RM4 then do;
237 WOE_RM4 = -1.240544367;
238 end;
239 end;
240
241 *-----*;
242 * Variable: RM5;
243 *-----*;
244 LABEL WOE_RM5 =
245 "Weight of Evidence: FactoringtoTotalRiskT0";
246
247 if MISSING(RM5) then do;
248 WOE_RM5 = -0.10749168;
249 end;
250 else if NOT MISSING(RM5) then do;
251 if RM5 < 0.65 then do;
252 WOE_RM5 = -1.274016667;
253 end;
254 else
255 if 0.65 <= RM5 then do;
256 WOE_RM5 = 0.3126087812;
257 end;
258 end;
259
260 *-----*;
261 * Variable: RM6;
262 *-----*;
263 LABEL WOE_RM6 =
264 "Weight of Evidence: TtlLimitUtilizationatRM6exzucT6";
265
266 if MISSING(RM6) then do;
267 WOE_RM6 = -0.066460066;
268 end;
269 else if NOT MISSING(RM6) then do;
270 if RM6 < 0.16 then do;
271 WOE_RM6 = 1.3180054234;
272 end;
273 else
274 if 0.16 <= RM6 AND RM6 < 0.35 then do;
275 WOE_RM6 = 0.510961557;
276 end;
277 else
278 if 0.35 <= RM6 AND RM6 < 0.57 then do;
279 WOE_RM6 = -0.217419704;
280 end;

```

Optimized SAS Code

```

281     else
282     if 0.57 <= RM6 then do;
283     WOE_RM6 = -0.686632659;
284     end;
285     end;
286
287     *-----*;
288     * Variable: RM8;
289     *-----*;
290     LABEL WOE_RM8 =
291     "Weight of Evidence: ShortTermCashtoTotalLimitT0";
292
293     if MISSING(RM8) then do;
294     WOE_RM8 = -0.10749168;
295     end;
296     else if NOT MISSING(RM8) then do;
297     if RM8 < 0.02 then do;
298     WOE_RM8 = 1.2565820439;
299     end;
300     else
301     if 0.02 <= RM8 AND RM8 < 0.15 then do;
302     WOE_RM8 = 0.2834829193;
303     end;
304     else
305     if 0.15 <= RM8 AND RM8 < 0.22 then do;
306     WOE_RM8 = -0.310396425;
307     end;
308     else
309     if 0.22 <= RM8 AND RM8 < 0.4 then do;
310     WOE_RM8 = -0.677334365;
311     end;
312     else
313     if 0.4 <= RM8 then do;
314     WOE_RM8 = -0.819687569;
315     end;
316     end;
317     *-----*;
318     * TOOL: Scorecard;
319     * TYPE: MODEL;
320     * NODE: Scorecard;
321     *-----*;
322     *****;
323     *** begin scoring code for regression;
324     *****;
325
326     length _WARN_ $4;
327     label _WARN_ = 'Warnings' ;
328
329     length I_GoodBad $ 12;
330     label I_GoodBad = 'Into: GoodBad' ;
331     *** Target Values;
332     array SCORECARDDRF [2] $12 _temporary_ ('1' '0' );
333     label U_GoodBad = 'Unnormalized Into: GoodBad' ;
334     *** Unnormalized target values;
335     ARRAY SCORECARDDRU[2] _TEMPORARY_ (1 0);
336
337     drop _DM_BAD;
338     _DM_BAD=0;
339
340     *** Check WOE_CEK1 for missing values ;
341     if missing( WOE_CEK1 ) then do;
342         substr(_warn_,1,1) = 'M';
343         _DM_BAD = 1;
344     end;
345
346     *** Check WOE_FINANSAL2 for missing values ;
347     if missing( WOE_FINANSAL2 ) then do;
348         substr(_warn_,1,1) = 'M';
349         _DM_BAD = 1;
350     end;

```

Optimized SAS Code

```

351
352   *** Check WOE_GECIKME1 for missing values ;
353   if missing( WOE_GECIKME1 ) then do;
354     substr(_warn_,1,1) = 'M';
355     _DM_BAD = 1;
356   end;
357
358   *** Check WOE_GECIKME2 for missing values ;
359   if missing( WOE_GECIKME2 ) then do;
360     substr(_warn_,1,1) = 'M';
361     _DM_BAD = 1;
362   end;
363
364   *** Check WOE_LIMIT3 for missing values ;
365   if missing( WOE_LIMIT3 ) then do;
366     substr(_warn_,1,1) = 'M';
367     _DM_BAD = 1;
368   end;
369
370   *** Check WOE_ORTAK for missing values ;
371   if missing( WOE_ORTAK ) then do;
372     substr(_warn_,1,1) = 'M';
373     _DM_BAD = 1;
374   end;
375
376   *** Check WOE_RM1 for missing values ;
377   if missing( WOE_RM1 ) then do;
378     substr(_warn_,1,1) = 'M';
379     _DM_BAD = 1;
380   end;
381
382   *** Check WOE_RM4 for missing values ;
383   if missing( WOE_RM4 ) then do;
384     substr(_warn_,1,1) = 'M';
385     _DM_BAD = 1;
386   end;
387
388   *** Check WOE_RM5 for missing values ;
389   if missing( WOE_RM5 ) then do;
390     substr(_warn_,1,1) = 'M';
391     _DM_BAD = 1;
392   end;
393
394   *** Check WOE_RM6 for missing values ;
395   if missing( WOE_RM6 ) then do;
396     substr(_warn_,1,1) = 'M';
397     _DM_BAD = 1;
398   end;
399
400   *** Check WOE_RM8 for missing values ;
401   if missing( WOE_RM8 ) then do;
402     substr(_warn_,1,1) = 'M';
403     _DM_BAD = 1;
404   end;
405   *** If missing inputs, use averages;
406   if _DM_BAD > 0 then do;
407     _PO = 0.0294102949;
408     _P1 = 0.9705097051;
409     goto SCORECARDRI;
410   end;
411
412   *** Compute Linear Predictor;
413   drop _TEMP;
414   drop _LPO;
415   _LPO = 0;
416
417   *** Effect: WOE_CEK1 ;
418   _TEMP = WOE_CEK1 ;
419   _LPO = _LPO + ( -0.20420066322992 * _TEMP);
420

```

Optimized SAS Code

```

421 *** Effect: WOE_FINANSAL2 ;
422 _TEMP = WOE_FINANSAL2 ;
423 _LPO = _LPO + ( -0.26196344011261 * _TEMP);
424
425 *** Effect: WOE_GECIKME1 ;
426 _TEMP = WOE_GECIKME1 ;
427 _LPO = _LPO + ( -0.37797490551211 * _TEMP);
428
429 *** Effect: WOE_GECIKME2 ;
430 _TEMP = WOE_GECIKME2 ;
431 _LPO = _LPO + ( -0.55163928704073 * _TEMP);
432
433 *** Effect: WOE_LIMIT3 ;
434 _TEMP = WOE_LIMIT3 ;
435 _LPO = _LPO + ( -0.44179218466457 * _TEMP);
436
437 *** Effect: WOE_ORTAK ;
438 _TEMP = WOE_ORTAK ;
439 _LPO = _LPO + ( -0.52638679141678 * _TEMP);
440
441 *** Effect: WOE_RM1 ;
442 _TEMP = WOE_RM1 ;
443 _LPO = _LPO + ( -0.47047241030407 * _TEMP);
444
445 *** Effect: WOE_RM4 ;
446 _TEMP = WOE_RM4 ;
447 _LPO = _LPO + ( -0.45406886893196 * _TEMP);
448
449 *** Effect: WOE_RM5 ;
450 _TEMP = WOE_RM5 ;
451 _LPO = _LPO + ( -0.39425051765572 * _TEMP);
452
453 *** Effect: WOE_RM6 ;
454 _TEMP = WOE_RM6 ;
455 _LPO = _LPO + ( -0.340459999144392 * _TEMP);
456
457 *** Effect: WOE_RM8 ;
458 _TEMP = WOE_RM8 ;
459 _LPO = _LPO + ( -0.27579561592594 * _TEMP);
460
461 *** Naive Posterior Probabilities;
462 drop _MAXP _IY _PO _P1;
463 _TEMP = -3.50243967500147 + _LPO;
464 if (_TEMP < 0) then do;
465     _TEMP = exp(_TEMP);
466     _PO = _TEMP / (1 + _TEMP);
467 end;
468 else _PO = 1 / (1 + exp(-_TEMP));
469 _P1 = 1.0 - _PO;
470
471 SCORECARDRI:
472
473
474 *** Posterior Probabilities and Predicted Level;
475 label P_GoodBad1 = 'Predicted: GoodBad=1' ;
476 label P_GoodBad0 = 'Predicted: GoodBad=0' ;
477 P_GoodBad1 = _PO;
478 _MAXP = _PO;
479 _IY = 1;
480 P_GoodBad0 = _P1;
481 if (_P1 > _MAXP + 1E-8) then do;
482     _MAXP = _P1;
483     _IY = 2;
484 end;
485 I_GoodBad = SCORECARDRI[_IY];
486 U_GoodBad = SCORECARDRI[_IY];
487
488 *****;
489 ***** end scoring code for regression;
490 *****;

```

Optimized SAS Code

```

491 *-----*
492 * generateScorepoints_note;
493 *-----*
494 SCORECARD_POINTS = 0;
495
496 *-----*
497 * Variable: CEK1;
498 *-----*
499 if MISSING(CEK1) then do;
500 SCORECARD_POINTS = SCORECARD_POINTS + 19;
501 SCR_CEK1= 19;
502 end;
503 else if NOT MISSING(CEK1) AND CEK1 < 7777 then do;
504 SCORECARD_POINTS = SCORECARD_POINTS + 12;
505 SCR_CEK1 = 12;
506 end;
507 else if NOT MISSING(CEK1) and 7777 <= CEK1 AND CEK1 < 9999 then do;
508 SCORECARD_POINTS = SCORECARD_POINTS + 15;
509 SCR_CEK1 = 15;
510 end;
511 else if NOT MISSING(CEK1) and 9999 <= CEK1 then do;
512 SCORECARD_POINTS = SCORECARD_POINTS + 19;
513 SCR_CEK1 = 19;
514 end;
515
516 *-----*
517 * Variable: FINANSAL2;
518 *-----*
519 if MISSING(FINANSAL2) then do;
520 SCORECARD_POINTS = SCORECARD_POINTS + 25;
521 SCR_FINANSAL2= 25;
522 end;
523 else if NOT MISSING(FINANSAL2) AND FINANSAL2 < 0.38 then do;
524 SCORECARD_POINTS = SCORECARD_POINTS + 13;
525 SCR_FINANSAL2 = 13;
526 end;
527 else if NOT MISSING(FINANSAL2) and 0.38 <= FINANSAL2 AND FINANSAL2 < 0.93 then do;
528 SCORECARD_POINTS = SCORECARD_POINTS + 18;
529 SCR_FINANSAL2 = 18;
530 end;
531 else if NOT MISSING(FINANSAL2) and 0.93 <= FINANSAL2 AND FINANSAL2 < 1.79 then do;
532 SCORECARD_POINTS = SCORECARD_POINTS + 21;
533 SCR_FINANSAL2 = 21;
534 end;
535 else if NOT MISSING(FINANSAL2) and 1.79 <= FINANSAL2 then do;
536 SCORECARD_POINTS = SCORECARD_POINTS + 23;
537 SCR_FINANSAL2 = 23;
538 end;
539
540 *-----*
541 * Variable: GECIKME1;
542 *-----*
543 if MISSING(GECIKME1) then do;
544 SCORECARD_POINTS = SCORECARD_POINTS + 25;
545 SCR_GECIKME1= 25;
546 end;
547 else if NOT MISSING(GECIKME1) AND GECIKME1 < 1 then do;
548 SCORECARD_POINTS = SCORECARD_POINTS + 25;
549 SCR_GECIKME1 = 25;
550 end;
551 else if NOT MISSING(GECIKME1) and 1 <= GECIKME1 AND GECIKME1 < 3 then do;
552 SCORECARD_POINTS = SCORECARD_POINTS + 14;
553 SCR_GECIKME1 = 14;
554 end;
555 else if NOT MISSING(GECIKME1) and 3 <= GECIKME1 AND GECIKME1 < 5 then do;
556 SCORECARD_POINTS = SCORECARD_POINTS + 11;
557 SCR_GECIKME1 = 11;
558 end;
559 else if NOT MISSING(GECIKME1) and 5 <= GECIKME1 AND GECIKME1 < 16 then do;
560 SCORECARD_POINTS = SCORECARD_POINTS + 6;

```

Optimized SAS Code

```

561 SCR_GECIKME1 = 6;
562 end;
563 else if NOT MISSING(GECIKME1) and 16 <= GECIKME1 then do;
564 SCORECARD_POINTS = SCORECARD_POINTS + 0;
565 SCR_GECIKME1 = 0;
566 end;
567
568 *-----*;
569 * Variable: GECIKME2;
570 *-----*;
571 if MISSING(GECIKME2) then do;
572 SCORECARD_POINTS = SCORECARD_POINTS + 21;
573 SCR_GECIKME2= 21;
574 end;
575 else if NOT MISSING(GECIKME2) AND GECIKME2 < 1 then do;
576 SCORECARD_POINTS = SCORECARD_POINTS + 21;
577 SCR_GECIKME2 = 21;
578 end;
579 else if NOT MISSING(GECIKME2) and 1 <= GECIKME2 AND GECIKME2 < 5 then do;
580 SCORECARD_POINTS = SCORECARD_POINTS + 9;
581 SCR_GECIKME2 = 9;
582 end;
583 else if NOT MISSING(GECIKME2) and 5 <= GECIKME2 AND GECIKME2 < 21 then do;
584 SCORECARD_POINTS = SCORECARD_POINTS + -2;
585 SCR_GECIKME2 = -2;
586 end;
587 else if NOT MISSING(GECIKME2) and 21 <= GECIKME2 then do;
588 SCORECARD_POINTS = SCORECARD_POINTS + -11;
589 SCR_GECIKME2 = -11;
590 end;
591
592 *-----*;
593 * Variable: LIMIT3;
594 *-----*;
595 if MISSING(LIMIT3) then do;
596 SCORECARD_POINTS = SCORECARD_POINTS + 25;
597 SCR_LIMIT3= 25;
598 end;
599 else if NOT MISSING(LIMIT3) AND LIMIT3 < 0.02 then do;
600 SCORECARD_POINTS = SCORECARD_POINTS + 31;
601 SCR_LIMIT3 = 31;
602 end;
603 else if NOT MISSING(LIMIT3) and 0.02 <= LIMIT3 AND LIMIT3 < 0.46 then do;
604 SCORECARD_POINTS = SCORECARD_POINTS + 17;
605 SCR_LIMIT3 = 17;
606 end;
607 else if NOT MISSING(LIMIT3) and 0.46 <= LIMIT3 AND LIMIT3 < 0.75 then do;
608 SCORECARD_POINTS = SCORECARD_POINTS + 10;
609 SCR_LIMIT3 = 10;
610 end;
611 else if NOT MISSING(LIMIT3) and 0.75 <= LIMIT3 then do;
612 SCORECARD_POINTS = SCORECARD_POINTS + 5;
613 SCR_LIMIT3 = 5;
614 end;
615
616 *-----*;
617 * Variable: ORTAK;
618 *-----*;
619 if MISSING(ORTAK) then do;
620 SCORECARD_POINTS = SCORECARD_POINTS + 28;
621 SCR_ORTAK= 28;
622 end;
623 else if NOT MISSING(ORTAK) AND ORTAK < 1 then do;
624 SCORECARD_POINTS = SCORECARD_POINTS + 23;
625 SCR_ORTAK = 23;
626 end;
627 else if NOT MISSING(ORTAK) and 1 <= ORTAK AND ORTAK < 2 then do;
628 SCORECARD_POINTS = SCORECARD_POINTS + 17;
629 SCR_ORTAK = 17;
630 end;

```

Optimized SAS Code

```

631 else if NOT MISSING(ORTAK) and 2 <= ORTAK AND ORTAK < 3 then do;
632 SCORECARD_POINTS = SCORECARD_POINTS + 11;
633 SCR_ORTAK = 11;
634 end;
635 else if NOT MISSING(ORTAK) and 3 <= ORTAK AND ORTAK < 4 then do;
636 SCORECARD_POINTS = SCORECARD_POINTS + 6;
637 SCR_ORTAK = 6;
638 end;
639 else if NOT MISSING(ORTAK) and 4 <= ORTAK then do;
640 SCORECARD_POINTS = SCORECARD_POINTS + 1;
641 SCR_ORTAK = 1;
642 end;
643
644 *-----*;
645 * Variable: RM1;
646 *-----*;
647 if MISSING(RM1) then do;
648 SCORECARD_POINTS = SCORECARD_POINTS + 16;
649 SCR_RM1= 16;
650 end;
651 else if NOT MISSING(RM1) AND RM1 < 2222 then do;
652 SCORECARD_POINTS = SCORECARD_POINTS + 0;
653 SCR_RM1 = 0;
654 end;
655 else if NOT MISSING(RM1) and 2222 <= RM1 AND RM1 < 6666 then do;
656 SCORECARD_POINTS = SCORECARD_POINTS + 5;
657 SCR_RM1 = 5;
658 end;
659 else if NOT MISSING(RM1) and 6666 <= RM1 then do;
660 SCORECARD_POINTS = SCORECARD_POINTS + 24;
661 SCR_RM1 = 24;
662 end;
663
664 *-----*;
665 * Variable: RM4;
666 *-----*;
667 if MISSING(RM4) then do;
668 SCORECARD_POINTS = SCORECARD_POINTS + 15;
669 SCR_RM4= 15;
670 end;
671 else if NOT MISSING(RM4) AND RM4 < 8 then do;
672 SCORECARD_POINTS = SCORECARD_POINTS + 25;
673 SCR_RM4 = 25;
674 end;
675 else if NOT MISSING(RM4) and 8 <= RM4 AND RM4 < 13 then do;
676 SCORECARD_POINTS = SCORECARD_POINTS + 18;
677 SCR_RM4 = 18;
678 end;
679 else if NOT MISSING(RM4) and 13 <= RM4 AND RM4 < 18 then do;
680 SCORECARD_POINTS = SCORECARD_POINTS + 10;
681 SCR_RM4 = 10;
682 end;
683 else if NOT MISSING(RM4) and 18 <= RM4 then do;
684 SCORECARD_POINTS = SCORECARD_POINTS + 1;
685 SCR_RM4 = 1;
686 end;
687
688 *-----*;
689 * Variable: RM5;
690 *-----*;
691 if MISSING(RM5) then do;
692 SCORECARD_POINTS = SCORECARD_POINTS + 16;
693 SCR_RM5= 16;
694 end;
695 else if NOT MISSING(RM5) AND RM5 < 0.65 then do;
696 SCORECARD_POINTS = SCORECARD_POINTS + 3;
697 SCR_RM5 = 3;
698 end;
699 else if NOT MISSING(RM5) and 0.65 <= RM5 then do;
700 SCORECARD_POINTS = SCORECARD_POINTS + 21;

```

Optimized SAS Code

```

701 SCR_RM5 = 21;
702 end;
703
704 *-----*;
705 * Variable: RM6;
706 *-----*;
707 if MISSING(RM6) then do;
708 SCORECARD_POINTS = SCORECARD_POINTS + 16;
709 SCR_RM6= 16;
710 end;
711 else if NOT MISSING(RM6) AND RM6 < 0.16 then do;
712 SCORECARD_POINTS = SCORECARD_POINTS + 30;
713 SCR_RM6 = 30;
714 end;
715 else if NOT MISSING(RM6) and 0.16 <= RM6 AND RM6 < 0.35 then do;
716 SCORECARD_POINTS = SCORECARD_POINTS + 22;
717 SCR_RM6 = 22;
718 end;
719 else if NOT MISSING(RM6) and 0.35 <= RM6 AND RM6 < 0.57 then do;
720 SCORECARD_POINTS = SCORECARD_POINTS + 15;
721 SCR_RM6 = 15;
722 end;
723 else if NOT MISSING(RM6) and 0.57 <= RM6 then do;
724 SCORECARD_POINTS = SCORECARD_POINTS + 10;
725 SCR_RM6 = 10;
726 end;
727
728 *-----*;
729 * Variable: RM8;
730 *-----*;
731 if MISSING(RM8) then do;
732 SCORECARD_POINTS = SCORECARD_POINTS + 16;
733 SCR_RM8= 16;
734 end;
735 else if NOT MISSING(RM8) AND RM8 < 0.02 then do;
736 SCORECARD_POINTS = SCORECARD_POINTS + 27;
737 SCR_RM8 = 27;
738 end;
739 else if NOT MISSING(RM8) and 0.02 <= RM8 AND RM8 < 0.15 then do;
740 SCORECARD_POINTS = SCORECARD_POINTS + 19;
741 SCR_RM8 = 19;
742 end;
743 else if NOT MISSING(RM8) and 0.15 <= RM8 AND RM8 < 0.22 then do;
744 SCORECARD_POINTS = SCORECARD_POINTS + 15;
745 SCR_RM8 = 15;
746 end;
747 else if NOT MISSING(RM8) and 0.22 <= RM8 AND RM8 < 0.4 then do;
748 SCORECARD_POINTS = SCORECARD_POINTS + 12;
749 SCR_RM8 = 12;
750 end;
751 else if NOT MISSING(RM8) and 0.4 <= RM8 then do;
752 SCORECARD_POINTS = SCORECARD_POINTS + 11;
753 SCR_RM8 = 11;
754 end;
755 *;
756 * Assign SCORECARD_BIN values;
757 *;
758 if SCORECARD_POINTS < 75 then SCORECARD_BIN = 1;
759 else if SCORECARD_POINTS < 83 then SCORECARD_BIN = 2;
760 else if SCORECARD_POINTS < 91 then SCORECARD_BIN = 3;
761 else if SCORECARD_POINTS < 99 then SCORECARD_BIN = 4;
762 else if SCORECARD_POINTS < 107 then SCORECARD_BIN = 5;
763 else if SCORECARD_POINTS < 115 then SCORECARD_BIN = 6;
764 else if SCORECARD_POINTS < 123 then SCORECARD_BIN = 7;
765 else if SCORECARD_POINTS < 131 then SCORECARD_BIN = 8;
766 else if SCORECARD_POINTS < 139 then SCORECARD_BIN = 9;
767 else if SCORECARD_POINTS < 147 then SCORECARD_BIN = 10;
768 else if SCORECARD_POINTS < 155 then SCORECARD_BIN = 11;
769 else if SCORECARD_POINTS < 163 then SCORECARD_BIN = 12;
770 else if SCORECARD_POINTS < 171 then SCORECARD_BIN = 13;

```


Optimized SAS Code

```

771 else if SCORECARD_POINTS < 179 then SCORECARD_BIN = 14;
772 else if SCORECARD_POINTS < 187 then SCORECARD_BIN = 15;
773 else if SCORECARD_POINTS < 195 then SCORECARD_BIN = 16;
774 else if SCORECARD_POINTS < 203 then SCORECARD_BIN = 17;
775 else if SCORECARD_POINTS < 211 then SCORECARD_BIN = 18;
776 else if SCORECARD_POINTS < 219 then SCORECARD_BIN = 19;
777 else if SCORECARD_POINTS < 227 then SCORECARD_BIN = 20;
778 else if SCORECARD_POINTS < 235 then SCORECARD_BIN = 21;
779 else if SCORECARD_POINTS < 243 then SCORECARD_BIN = 22;
780 else if SCORECARD_POINTS < 251 then SCORECARD_BIN = 23;
781 else if SCORECARD_POINTS < 259 then SCORECARD_BIN = 24;
782 else SCORECARD_BIN = 25;
783 *-----*;
784 * TOOL: Score Node;
785 * TYPE: ASSESS;
786 * NODE: Score;
787 *-----*;
788 *-----*;
789 * Score: Creating Fixed Names;
790 *-----*;
791 LABEL EM_SEGMENT = 'Segment';
792 EM_SEGMENT = SCORECARD_BIN;
793 LABEL EM_EVENTPROBABILITY = 'Probability for level 1 of GoodBad';
794 EM_EVENTPROBABILITY = P_GoodBad1;
795 LABEL EM_PROBABILITY = 'Probability of Classification';
796 EM_PROBABILITY =
797 max(
798 P_GoodBad1
799 ,
800 P_GoodBad0
801 );
802 LENGTH EM_CLASSIFICATION $4096;
803 LABEL EM_CLASSIFICATION = "Prediction for GoodBad";
804 EM_CLASSIFICATION = I_GoodBad;
805

```

Ek C. Cox Regresyon Analizi SAS Çıktıları

```
Optimized SAS Code
1  *-----*;
2  * EM SCORE CODE;
3  *-----*;
4  *-----*;
5  * TOOL: Input Data Source;
6  * TYPE: SAMPLE;
7  * NODE: Ids2;
8  *-----*;
9  *-----*;
10 * TOOL: Partition Class;
11 * TYPE: SAMPLE;
12 * NODE: Part2;
13 *-----*;
14 *-----*;
15 * TOOL: Extension Class;
16 * TYPE: CREDSCORE;
17 * NODE: IGM2;
18 *-----*;
19 length _UFormat $200;
20 drop _UFormat;
21 _UFormat='';
22
23 *-----*;
24 * Variable: CEK1;
25 *-----*;
26 LABEL WOE_CEK1 = "Weight of Evidence: CEK1";
27
28 if MISSING(CEK1) then do;
29   WOE_CEK1 = 0.3668207366;
30 end;
31 else if NOT MISSING(CEK1) then do;
32   if CEK1 < 7777 then do;
33     WOE_CEK1 = -1.153569765;
34   end;
35   else
36     if 7777 <= CEK1 AND CEK1 < 9999 then do;
37       WOE_CEK1 = -0.354935471;
38     end;
39     else
40       if 9999 <= CEK1 then do;
41         WOE_CEK1 = 0.3668207366;
42       end;
43     end;
44
45 *-----*;
46 * Variable: FINANSAL2;
47 *-----*;
48 LABEL WOE_FINANSAL2 = "Weight of Evidence: FINANSAL2";
49
50 if MISSING(FINANSAL2) then do;
51   WOE_FINANSAL2 = -0.429086178;
52 end;
53 else if NOT MISSING(FINANSAL2) then do;
54   if FINANSAL2 < 0.28 then do;
55     WOE_FINANSAL2 = -0.429086178;
56   end;
57   else
58     if 0.28 <= FINANSAL2 AND FINANSAL2 < 0.4 then do;
59       WOE_FINANSAL2 = -0.308243511;
60     end;
61     else
62       if 0.4 <= FINANSAL2 AND FINANSAL2 < 0.82 then do;
63         WOE_FINANSAL2 = 0.0463287305;
64       end;
65       else
66         if 0.82 <= FINANSAL2 then do;
67           WOE_FINANSAL2 = 0.5826981342;
68         end;
69       end;
70
```

Optimized SAS Code

```

71  *-----*
72  * Variable: GECIKME1;
73  *-----*
74  LABEL WOE_GECIKME1 = "Weight of Evidence: GECIKME1";
75
76  if MISSING(GECIKME1) then do;
77  WOE_GECIKME1 = 0.5551859059;
78  end;
79  else if NOT MISSING(GECIKME1) then do;
80  if GECIKME1 < 2 then do;
81  WOE_GECIKME1 = 0.5551859059;
82  end;
83  else
84  if 2 <= GECIKME1 AND GECIKME1 < 7 then do;
85  WOE_GECIKME1 = -0.596564765;
86  end;
87  else
88  if 7 <= GECIKME1 AND GECIKME1 < 15 then do;
89  WOE_GECIKME1 = -1.002029873;
90  end;
91  else
92  if 15 <= GECIKME1 AND GECIKME1 < 35 then do;
93  WOE_GECIKME1 = -1.106726819;
94  end;
95  else
96  if 35 <= GECIKME1 then do;
97  WOE_GECIKME1 = -1.807752856;
98  end;
99  end;
100
101  *-----*
102  * Variable: LIMIT1;
103  *-----*
104  LABEL WOE_LIMIT1 = "Weight of Evidence: LIMIT1";
105
106  if MISSING(LIMIT1) then do;
107  WOE_LIMIT1 = 0.7409140356;
108  end;
109  else if NOT MISSING(LIMIT1) then do;
110  if LIMIT1 < 100000 then do;
111  WOE_LIMIT1 = 0.7409140356;
112  end;
113  else
114  if 100000 <= LIMIT1 AND LIMIT1 < 750000 then do;
115  WOE_LIMIT1 = -0.422257336;
116  end;
117  else
118  if 750000 <= LIMIT1 AND LIMIT1 < 1000000 then do;
119  WOE_LIMIT1 = -0.048629131;
120  end;
121  else
122  if 1000000 <= LIMIT1 then do;
123  WOE_LIMIT1 = 0.4509722239;
124  end;
125  end;
126
127  *-----*
128  * Variable: LIMIT2;
129  *-----*
130  LABEL WOE_LIMIT2 = "Weight of Evidence: LIMIT2";
131
132  if MISSING(LIMIT2) then do;
133  WOE_LIMIT2 = 1.0060240772;
134  end;
135  else if NOT MISSING(LIMIT2) then do;
136  if LIMIT2 < 123 then do;
137  WOE_LIMIT2 = -0.940228423;
138  end;
139  else
140  if 123 <= LIMIT2 AND LIMIT2 < 237 then do;

```

Optimized SAS Code

```

141 WOE_LIMIT2 = -0.026432009;
142 end;
143 else
144 if 237 <= LIMIT2 AND LIMIT2 < 367 then do;
145 WOE_LIMIT2 = -0.753825867;
146 end;
147 else
148 if 367 <= LIMIT2 then do;
149 WOE_LIMIT2 = -0.052741223;
150 end;
151 end;
152
153 *-----*;
154 * Variable: LIMIT3;
155 *-----*;
156 LABEL WOE_LIMIT3 = "Weight of Evidence: LIMIT3";
157
158 if MISSING(LIMIT3) then do;
159 WOE_LIMIT3 = 0.9854517286;
160 end;
161 else if NOT MISSING(LIMIT3) then do;
162 if LIMIT3 < 0.04 then do;
163 WOE_LIMIT3 = 0.9854517286;
164 end;
165 else
166 if 0.04 <= LIMIT3 AND LIMIT3 < 0.25 then do;
167 WOE_LIMIT3 = 0.1428334334;
168 end;
169 else
170 if 0.25 <= LIMIT3 AND LIMIT3 < 0.42 then do;
171 WOE_LIMIT3 = -0.366331779;
172 end;
173 else
174 if 0.42 <= LIMIT3 AND LIMIT3 < 0.83 then do;
175 WOE_LIMIT3 = -0.654665876;
176 end;
177 else
178 if 0.83 <= LIMIT3 then do;
179 WOE_LIMIT3 = -1.260635681;
180 end;
181 end;
182
183 *-----*;
184 * Variable: MEVDUAT;
185 *-----*;
186 LABEL WOE_MEVDUAT = "Weight of Evidence: MEVDUAT";
187
188 if MISSING(MEVDUAT) then do;
189 WOE_MEVDUAT = -0.85617661;
190 end;
191 else if NOT MISSING(MEVDUAT) then do;
192 if MEVDUAT < 101.56 then do;
193 WOE_MEVDUAT = 0.3124137514;
194 end;
195 else
196 if 101.56 <= MEVDUAT AND MEVDUAT < 6154.13 then do;
197 WOE_MEVDUAT = -0.124357969;
198 end;
199 else
200 if 6154.13 <= MEVDUAT then do;
201 WOE_MEVDUAT = 0.6274558807;
202 end;
203 end;
204
205 *-----*;
206 * Variable: ORTAK;
207 *-----*;
208 LABEL WOE_ORTAK = "Weight of Evidence: ORTAK";
209
210 if MISSING(ORTAK) then do;

```

Optimized SAS Code

```

211   WOE_OR TAK = 0.4797120473;
212   end;
213   else if NOT MISSING(OR TAK) then do;
214   if OR TAK < 1 then do;
215   WOE_OR TAK = 0.4797120473;
216   end;
217   else
218   if 1 <= OR TAK AND OR TAK < 2 then do;
219   WOE_OR TAK = -0.046265269;
220   end;
221   else
222   if 2 <= OR TAK AND OR TAK < 3 then do;
223   WOE_OR TAK = -0.388580394;
224   end;
225   else
226   if 3 <= OR TAK then do;
227   WOE_OR TAK = -0.73054873;
228   end;
229   end;
230
231   *-----*
232   * Variable: RM1;
233   *-----*
234   LABEL WOE_RM1 = "Weight of Evidence: RM1";
235
236   if MISSING(RM1) then do;
237   WOE_RM1 = 0.4428722915;
238   end;
239   else if NOT MISSING(RM1) then do;
240   if RM1 < 7.08 then do;
241   WOE_RM1 = -1.141544321;
242   end;
243   else
244   if 7.08 <= RM1 AND RM1 < 6666 then do;
245   WOE_RM1 = -0.984616378;
246   end;
247   else
248   if 6666 <= RM1 then do;
249   WOE_RM1 = 0.4428722915;
250   end;
251   end;
252
253   *-----*
254   * Variable: RM2;
255   *-----*
256   LABEL WOE_RM2 = "Weight of Evidence: RM2";
257
258   if MISSING(RM2) then do;
259   WOE_RM2 = 0.2914224664;
260   end;
261   else if NOT MISSING(RM2) then do;
262   if RM2 < 8888 then do;
263   WOE_RM2 = -0.699976234;
264   end;
265   else
266   if 8888 <= RM2 AND RM2 < 9999 then do;
267   WOE_RM2 = -1.662749238;
268   end;
269   else
270   if 9999 <= RM2 then do;
271   WOE_RM2 = 0.2914224664;
272   end;
273   end;
274
275   *-----*
276   * Variable: RM6;
277   *-----*
278   LABEL WOE_RM6 = "Weight of Evidence: RM6";
279
280   if MISSING(RM6) then do;

```

Optimized SAS Code

```

281 WOE_RM6 = -0.134502132;
282 end;
283 else if NOT MISSING(RM6) then do;
284 if RM6 < 0.27 then do;
285 WOE_RM6 = 1.023600273;
286 end;
287 else
288 if 0.27 <= RM6 AND RM6 < 0.4 then do;
289 WOE_RM6 = 0.1822564769;
290 end;
291 else
292 if 0.4 <= RM6 AND RM6 < 0.48 then do;
293 WOE_RM6 = -0.201627749;
294 end;
295 else
296 if 0.48 <= RM6 AND RM6 < 0.57 then do;
297 WOE_RM6 = -0.471561481;
298 end;
299 else
300 if 0.57 <= RM6 then do;
301 WOE_RM6 = -0.746301852;
302 end;
303 end;
304
305 *-----*;
306 * Variable: RM8;
307 *-----*;
308 LABEL WOE_RM8 = "Weight of Evidence: RM8";
309
310 if MISSING(RM8) then do;
311 WOE_RM8 = 1.2023539823;
312 end;
313 else if NOT MISSING(RM8) then do;
314 if RM8 < 0.02 then do;
315 WOE_RM8 = 1.2023539823;
316 end;
317 else
318 if 0.02 <= RM8 AND RM8 < 0.08 then do;
319 WOE_RM8 = 0.6130442998;
320 end;
321 else
322 if 0.08 <= RM8 AND RM8 < 0.12 then do;
323 WOE_RM8 = 0.004274632;
324 end;
325 else
326 if 0.12 <= RM8 AND RM8 < 0.18 then do;
327 WOE_RM8 = -0.102342451;
328 end;
329 else
330 if 0.18 <= RM8 AND RM8 < 0.25 then do;
331 WOE_RM8 = -0.46989533;
332 end;
333 else
334 if 0.25 <= RM8 AND RM8 < 0.36 then do;
335 WOE_RM8 = -0.788083824;
336 end;
337 else
338 if 0.36 <= RM8 then do;
339 WOE_RM8 = -1.023307271;
340 end;
341 end;
342
343 *-----*;
344 * Variable: YAS2;
345 *-----*;
346 LABEL WOE_YAS2 = "Weight of Evidence: YAS2";
347
348 if MISSING(YAS2) then do;
349 WOE_YAS2 = 0.6499676539;
350 end;

```

Optimized SAS Code

```

351     else if NOT MISSING(YAS2) then do;
352     if YAS2 < 466 then do;
353     WOE_YAS2 = -0.626125709;
354     end;
355     else
356     if 466 <= YAS2 AND YAS2 < 962 then do;
357     WOE_YAS2 = -0.325713583;
358     end;
359     else
360     if 962 <= YAS2 AND YAS2 < 1230 then do;
361     WOE_YAS2 = -0.029502301;
362     end;
363     else
364     if 1230 <= YAS2 AND YAS2 < 1493 then do;
365     WOE_YAS2 = -0.064093427;
366     end;
367     else
368     if 1493 <= YAS2 AND YAS2 < 1819 then do;
369     WOE_YAS2 = 0.149092361;
370     end;
371     else
372     if 1819 <= YAS2 AND YAS2 < 2247 then do;
373     WOE_YAS2 = 0.0469519777;
374     end;
375     else
376     if 2247 <= YAS2 then do;
377     WOE_YAS2 = 0.6499676539;
378     end;
379     end;
380     *-----?;
381     * TOOL: Extension Class;
382     * TYPE: APPS;
383     * NODE: SURV3;
384     *-----?;
385     *****;
386     *** begin scoring code for regression;
387     *****;
388
389     length _WARN_ $4;
390     label _WARN_ = 'Warnings' ;
391
392     length I__g_ $ 12;
393     label I__g_ = 'Into: _g_ ' ;
394     *** Target Values;
395     array SURV3DRF [2] $12 _temporary_ ('1' '0' );
396     label U__g_ = 'Unnormalized Into: _g_ ' ;
397     *** Unnormalized target values;
398     ARRAY SURV3DRU[2] _TEMPORARY_ (1 0);
399
400     drop _DM_BAD;
401     _DM_BAD=0;
402
403     /*-----*/
404     /*Survival Score Code*/
405     /*-----*/
406
407     label EM_SURVIVAL = "Survival Probability at Censoring Time";
408     label EM_SURVFCST = "Survival Probability at Future Time";
409     label EM_SURVEVENT = "Event Probability before or at the Future Time";
410     label EM_HAZARD = "Hazard Function at Censoring Time";
411     label EM_HZRDFCST = "Hazard Function at Future Time";
412
413     BadObs = 0;
414     if _T_ ne . and BadObs=0 then do;
415
416     label T_FCST = "Number of Time Units in Future";
417     T_FCST=_T_+2 ;
418
419
420     /*-----Generate Cubic Spline Basis Functions-----*/

```

Optimized SAS Code

```

421 if _t_ > 1 then
422   _csb1 = (_t_-1)**3 - _t_**3 + 3*_t_**2 - 3*_t_;
423 else
424   _csb1 = -_t_**3 + 3*_t_**2 - 3*_t_;
425
426 if _t_ > 2 then
427   _csb2 = (_t_-2)**3 - _t_**3 + 6*_t_**2 - 12*_t_;
428 else
429   _csb2 = -_t_**3 + 6*_t_**2 - 12*_t_;
430
431 if _t_ > 3 then
432   _csb3 = (_t_-3)**3 - _t_**3 + 9*_t_**2 - 27*_t_;
433 else
434   _csb3 = -_t_**3 + 9*_t_**2 - 27*_t_;
435
436 if _t_ > 4 then
437   _csb4 = (_t_-4)**3 - _t_**3 + 12*_t_**2 - 48*_t_;
438 else
439   _csb4 = -_t_**3 + 12*_t_**2 - 48*_t_;
440
441 if _t_ > 5 then
442   _csb5 = (_t_-5)**3 - _t_**3 + 15*_t_**2 - 75*_t_;
443 else
444   _csb5 = -_t_**3 + 15*_t_**2 - 75*_t_;
445
446 *** Check _t_ for missing values ;
447 if missing( _t_ ) then do;
448   substr(_warn_,1,1) = 'M';
449   _DM_BAD = 1;
450 end;
451
452 *** Check _csb1 for missing values ;
453 if missing( _csb1 ) then do;
454   substr(_warn_,1,1) = 'M';
455   _DM_BAD = 1;
456 end;
457
458 *** Check _csb2 for missing values ;
459 if missing( _csb2 ) then do;
460   substr(_warn_,1,1) = 'M';
461   _DM_BAD = 1;
462 end;
463
464 *** Check _csb3 for missing values ;
465 if missing( _csb3 ) then do;
466   substr(_warn_,1,1) = 'M';
467   _DM_BAD = 1;
468 end;
469
470 *** Check _csb4 for missing values ;
471 if missing( _csb4 ) then do;
472   substr(_warn_,1,1) = 'M';
473   _DM_BAD = 1;
474 end;
475
476 *** Check _csb5 for missing values ;
477 if missing( _csb5 ) then do;
478   substr(_warn_,1,1) = 'M';
479   _DM_BAD = 1;
480 end;
481
482 *** Check WOE_CEK1 for missing values ;
483 if missing( WOE_CEK1 ) then do;
484   substr(_warn_,1,1) = 'M';
485   _DM_BAD = 1;
486 end;
487
488 *** Check WOE_FINANSAL2 for missing values ;
489 if missing( WOE_FINANSAL2 ) then do;
490   substr(_warn_,1,1) = 'M';

```


Optimized SAS Code

```

491     _DM_BAD = 1;
492 end;
493
494 *** Check WOE_GECIKME1 for missing values ;
495 if missing( WOE_GECIKME1 ) then do;
496     substr(_warn_,1,1) = 'M';
497     _DM_BAD = 1;
498 end;
499
500 *** Check WOE_LIMIT1 for missing values ;
501 if missing( WOE_LIMIT1 ) then do;
502     substr(_warn_,1,1) = 'M';
503     _DM_BAD = 1;
504 end;
505
506 *** Check WOE_LIMIT2 for missing values ;
507 if missing( WOE_LIMIT2 ) then do;
508     substr(_warn_,1,1) = 'M';
509     _DM_BAD = 1;
510 end;
511
512 *** Check WOE_LIMIT3 for missing values ;
513 if missing( WOE_LIMIT3 ) then do;
514     substr(_warn_,1,1) = 'M';
515     _DM_BAD = 1;
516 end;
517
518 *** Check WOE_MEVDUAT for missing values ;
519 if missing( WOE_MEVDUAT ) then do;
520     substr(_warn_,1,1) = 'M';
521     _DM_BAD = 1;
522 end;
523
524 *** Check WOE_ORTAK for missing values ;
525 if missing( WOE_ORTAK ) then do;
526     substr(_warn_,1,1) = 'M';
527     _DM_BAD = 1;
528 end;
529
530 *** Check WOE_RM1 for missing values ;
531 if missing( WOE_RM1 ) then do;
532     substr(_warn_,1,1) = 'M';
533     _DM_BAD = 1;
534 end;
535
536 *** Check WOE_RM2 for missing values ;
537 if missing( WOE_RM2 ) then do;
538     substr(_warn_,1,1) = 'M';
539     _DM_BAD = 1;
540 end;
541
542 *** Check WOE_RM6 for missing values ;
543 if missing( WOE_RM6 ) then do;
544     substr(_warn_,1,1) = 'M';
545     _DM_BAD = 1;
546 end;
547
548 *** Check WOE_RM8 for missing values ;
549 if missing( WOE_RM8 ) then do;
550     substr(_warn_,1,1) = 'M';
551     _DM_BAD = 1;
552 end;
553
554 *** Check WOE_YAS2 for missing values ;
555 if missing( WOE_YAS2 ) then do;
556     substr(_warn_,1,1) = 'M';
557     _DM_BAD = 1;
558 end;
559 *** If missing inputs, use averages;
560 if _DM_BAD > 0 then do;

```

Optimized SAS Code

```

561     _PO = 0.0148392096;
562     _PI = 0.9851607904;
563     goto SURV3DR1;
564 end;
565
566 *** Compute Linear Predictor;
567 drop _TEMP;
568 drop _LPO ;
569 _LPO = 0;
570
571 *** Effect: _t_ ;
572 _TEMP = _t_ ;
573 _LPO = _LPO + (          0.45236540027388 * _TEMP);
574
575 *** Effect: _csb1 ;
576 _TEMP = _csb1 ;
577 _LPO = _LPO + (          321.511860630403 * _TEMP);
578
579 *** Effect: _csb2 ;
580 _TEMP = _csb2 ;
581 _LPO = _LPO + (         -89.792995693235 * _TEMP);
582
583 *** Effect: _csb3 ;
584 _TEMP = _csb3 ;
585 _LPO = _LPO + (          23.6714604062465 * _TEMP);
586
587 *** Effect: _csb4 ;
588 _TEMP = _csb4 ;
589 _LPO = _LPO + (         -5.50690245834361 * _TEMP);
590
591 *** Effect: _csb5 ;
592 _TEMP = _csb5 ;
593 _LPO = _LPO + (          0.79628206048608 * _TEMP);
594
595 *** Effect: WOE_CEK1 ;
596 _TEMP = WOE_CEK1 ;
597 _LPO = _LPO + (         -0.15002905083594 * _TEMP);
598
599 *** Effect: WOE_FINANSAL2 ;
600 _TEMP = WOE_FINANSAL2 ;
601 _LPO = _LPO + (         -0.17919861991054 * _TEMP);
602
603 *** Effect: WOE_GECIKME1 ;
604 _TEMP = WOE_GECIKME1 ;
605 _LPO = _LPO + (         -0.18956662344206 * _TEMP);
606
607 *** Effect: WOE_LIMIT1 ;
608 _TEMP = WOE_LIMIT1 ;
609 _LPO = _LPO + (         -0.34851630927994 * _TEMP);
610
611 *** Effect: WOE_LIMIT2 ;
612 _TEMP = WOE_LIMIT2 ;
613 _LPO = _LPO + (         -0.85808385338674 * _TEMP);
614
615 *** Effect: WOE_LIMIT3 ;
616 _TEMP = WOE_LIMIT3 ;
617 _LPO = _LPO + (         -0.52628321847465 * _TEMP);
618
619 *** Effect: WOE_MEYDUAT ;
620 _TEMP = WOE_MEYDUAT ;
621 _LPO = _LPO + (         -0.88776356456425 * _TEMP);
622
623 *** Effect: WOE_ORTAK ;
624 _TEMP = WOE_ORTAK ;
625 _LPO = _LPO + (         -0.72033869754683 * _TEMP);
626
627 *** Effect: WOE_RMI ;
628 _TEMP = WOE_RMI ;
629 _LPO = _LPO + (         -0.36953980811752 * _TEMP);
630

```

Optimized SAS Code

```

631 *** Effect: WOE_RM2 ;
632 _TEMP = WOE_RM2 ;
633 _LPO = _LPO + ( -0.34433996514399 * _TEMP);
634
635 *** Effect: WOE_RM6 ;
636 _TEMP = WOE_RM6 ;
637 _LPO = _LPO + ( -0.2351286534026 * _TEMP);
638
639 *** Effect: WOE_RM8 ;
640 _TEMP = WOE_RM8 ;
641 _LPO = _LPO + ( -0.24817155145679 * _TEMP);
642
643 *** Effect: WOE_YAS2 ;
644 _TEMP = WOE_YAS2 ;
645 _LPO = _LPO + ( -1.75260650797525 * _TEMP);
646
647 *** Effect: _t_*WOE_GECIKME1 ;
648 _TEMP = _t_ * WOE_GECIKME1 ;
649 _LPO = _LPO + ( -0.07164557184268 * _TEMP);
650
651 *** Effect: _t_*WOE_ORTAK ;
652 _TEMP = _t_ * WOE_ORTAK ;
653 _LPO = _LPO + ( 0.08821668142176 * _TEMP);
654
655 *** Naive Posterior Probabilities;
656 drop _MAXP _IY _PO _P1;
657 drop _LPMAX;
658 _LPMAX= 0;
659 _LPO = -16.6541828432845 + _LPO;
660 if _LPMAX < _LPO then _LPMAX = _LPO;
661 _LPO = exp(_LPO - _LPMAX);
662 _LPMAX = exp(-_LPMAX);
663 _P1 = 1 / (_LPMAX + _LPO);
664 _PO = _LPO * _P1;
665 _P1 = _LPMAX * _P1;
666
667 SURV3DR1:
668
669
670 *** Posterior Probabilities and Predicted Level;
671 label P_g_1 = 'Predicted: _g_ =1' ;
672 label P_g_0 = 'Predicted: _g_ =0' ;
673 P_g_1 = _PO;
674 _MAXP = _PO;
675 _IY = 1;
676 P_g_0 = _P1;
677 if (_P1 > _MAXP + 1E-8) then do;
678 _MAXP = _P1;
679 _IY = 2;
680 end;
681 I_g_ = SURV3DRF[_IY];
682 U_g_ = SURV3DRU[_IY];
683
684 *****;
685 ***** end scoring code for regression;
686 *****;
687
688 EM_SUBHZRD1 = P_g_1;
689 EM_SUBHZRD0 =1-(EM_SUBHZRD1);
690 EM_HAZARD=EM_SUBHZRD1;
691
692 /*----- Survival Function Estimation -----*/
693 EM_SURVFCST=1;
694 _t0_ = _t_;
695 t0_fcst = t_fcst;
696 _t_ = 0;
697 do while (_t_ <= t0_fcst);
698
699
700 /*-----Generate Cubic Spline Basis Functions-----*/

```

Optimized SAS Code

```

701  if _T_ > 1 then
702  _csb1 = (_T_-1)**3 - _T_**3 + 3*_T_**2 - 3*_T_;
703  else
704  _csb1 = -_T_**3 + 3*_T_**2 - 3*_T_;
705
706  if _T_ > 2 then
707  _csb2 = (_T_-2)**3 - _T_**3 + 6*_T_**2 - 12*_T_;
708  else
709  _csb2 = -_T_**3 + 6*_T_**2 - 12*_T_;
710
711  if _T_ > 3 then
712  _csb3 = (_T_-3)**3 - _T_**3 + 9*_T_**2 - 27*_T_;
713  else
714  _csb3 = -_T_**3 + 9*_T_**2 - 27*_T_;
715
716  if _T_ > 4 then
717  _csb4 = (_T_-4)**3 - _T_**3 + 12*_T_**2 - 48*_T_;
718  else
719  _csb4 = -_T_**3 + 12*_T_**2 - 48*_T_;
720
721  if _T_ > 5 then
722  _csb5 = (_T_-5)**3 - _T_**3 + 15*_T_**2 - 75*_T_;
723  else
724  _csb5 = -_T_**3 + 15*_T_**2 - 75*_T_;
725
726  *** Check _t_ for missing values ;
727  if missing( _t_ ) then do;
728  substr(_warn_,1,1) = 'M';
729  _DM_BAD = 1;
730  end;
731
732  *** Check _csb1 for missing values ;
733  if missing( _csb1 ) then do;
734  substr(_warn_,1,1) = 'M';
735  _DM_BAD = 1;
736  end;
737
738  *** Check _csb2 for missing values ;
739  if missing( _csb2 ) then do;
740  substr(_warn_,1,1) = 'M';
741  _DM_BAD = 1;
742  end;
743
744  *** Check _csb3 for missing values ;
745  if missing( _csb3 ) then do;
746  substr(_warn_,1,1) = 'M';
747  _DM_BAD = 1;
748  end;
749
750  *** Check _csb4 for missing values ;
751  if missing( _csb4 ) then do;
752  substr(_warn_,1,1) = 'M';
753  _DM_BAD = 1;
754  end;
755
756  *** Check _csb5 for missing values ;
757  if missing( _csb5 ) then do;
758  substr(_warn_,1,1) = 'M';
759  _DM_BAD = 1;
760  end;
761
762  *** Check WOE_CER1 for missing values ;
763  if missing( WOE_CER1 ) then do;
764  substr(_warn_,1,1) = 'M';
765  _DM_BAD = 1;
766  end;
767
768  *** Check WOE_FINANSAL2 for missing values ;
769  if missing( WOE_FINANSAL2 ) then do;
770  substr(_warn_,1,1) = 'M';

```

Optimized SAS Code

```

771     _DM_BAD = 1;
772 end;
773
774 *** Check WOE_GECIKME1 for missing values ;
775 if missing( WOE_GECIKME1 ) then do;
776     substr(_warn_,1,1) = 'M';
777     _DM_BAD = 1;
778 end;
779
780 *** Check WOE_LIMIT1 for missing values ;
781 if missing( WOE_LIMIT1 ) then do;
782     substr(_warn_,1,1) = 'M';
783     _DM_BAD = 1;
784 end;
785
786 *** Check WOE_LIMIT2 for missing values ;
787 if missing( WOE_LIMIT2 ) then do;
788     substr(_warn_,1,1) = 'M';
789     _DM_BAD = 1;
790 end;
791
792 *** Check WOE_LIMIT3 for missing values ;
793 if missing( WOE_LIMIT3 ) then do;
794     substr(_warn_,1,1) = 'M';
795     _DM_BAD = 1;
796 end;
797
798 *** Check WOE_MEVDUAT for missing values ;
799 if missing( WOE_MEVDUAT ) then do;
800     substr(_warn_,1,1) = 'M';
801     _DM_BAD = 1;
802 end;
803
804 *** Check WOE_ORTAK for missing values ;
805 if missing( WOE_ORTAK ) then do;
806     substr(_warn_,1,1) = 'M';
807     _DM_BAD = 1;
808 end;
809
810 *** Check WOE_RM1 for missing values ;
811 if missing( WOE_RM1 ) then do;
812     substr(_warn_,1,1) = 'M';
813     _DM_BAD = 1;
814 end;
815
816 *** Check WOE_RM2 for missing values ;
817 if missing( WOE_RM2 ) then do;
818     substr(_warn_,1,1) = 'M';
819     _DM_BAD = 1;
820 end;
821
822 *** Check WOE_RM6 for missing values ;
823 if missing( WOE_RM6 ) then do;
824     substr(_warn_,1,1) = 'M';
825     _DM_BAD = 1;
826 end;
827
828 *** Check WOE_RM8 for missing values ;
829 if missing( WOE_RM8 ) then do;
830     substr(_warn_,1,1) = 'M';
831     _DM_BAD = 1;
832 end;
833
834 *** Check WOE_YAS2 for missing values ;
835 if missing( WOE_YAS2 ) then do;
836     substr(_warn_,1,1) = 'M';
837     _DM_BAD = 1;
838 end;
839 *** If missing inputs, use averages;
840 if _DM_BAD > 0 then do;

```

Optimized SAS Code

```

841     _PO = 0.0148392096;
842     _P1 = 0.9851607904;
843     goto SURV3DR4;
844 end;
845
846 *** Compute Linear Predictor;
847 drop _TEMP;
848 drop _LPO ;
849 _LPO = 0;
850
851 *** Effect: _t_ ;
852 _TEMP = _t_ ;
853 _LPO = _LPO + (           0.45236540027308 * _TEMP);
854
855 *** Effect: _csb1 ;
856 _TEMP = _csb1 ;
857 _LPO = _LPO + (           321.511860630403 * _TEMP);
858
859 *** Effect: _csb2 ;
860 _TEMP = _csb2 ;
861 _LPO = _LPO + (          -89.792995693235 * _TEMP);
862
863 *** Effect: _csb3 ;
864 _TEMP = _csb3 ;
865 _LPO = _LPO + (           23.6714604062465 * _TEMP);
866
867 *** Effect: _csb4 ;
868 _TEMP = _csb4 ;
869 _LPO = _LPO + (          -5.50690245834361 * _TEMP);
870
871 *** Effect: _csb5 ;
872 _TEMP = _csb5 ;
873 _LPO = _LPO + (           0.79628206048608 * _TEMP);
874
875 *** Effect: WOE_CEK1 ;
876 _TEMP = WOE_CEK1 ;
877 _LPO = _LPO + (          -0.15002905083594 * _TEMP);
878
879 *** Effect: WOE_FINANSAL2 ;
880 _TEMP = WOE_FINANSAL2 ;
881 _LPO = _LPO + (          -0.17919861991054 * _TEMP);
882
883 *** Effect: WOE_GECIKME1 ;
884 _TEMP = WOE_GECIKME1 ;
885 _LPO = _LPO + (          -0.18956662344206 * _TEMP);
886
887 *** Effect: WOE_LIMIT1 ;
888 _TEMP = WOE_LIMIT1 ;
889 _LPO = _LPO + (          -0.34851630927994 * _TEMP);
890
891 *** Effect: WOE_LIMIT2 ;
892 _TEMP = WOE_LIMIT2 ;
893 _LPO = _LPO + (          -0.85008385338674 * _TEMP);
894
895 *** Effect: WOE_LIMIT3 ;
896 _TEMP = WOE_LIMIT3 ;
897 _LPO = _LPO + (          -0.52628321847465 * _TEMP);
898
899 *** Effect: WOE_MEVDUAT ;
900 _TEMP = WOE_MEVDUAT ;
901 _LPO = _LPO + (          -0.88776356456425 * _TEMP);
902
903 *** Effect: WOE_ORTAK ;
904 _TEMP = WOE_ORTAK ;
905 _LPO = _LPO + (          -0.72033869754683 * _TEMP);
906
907 *** Effect: WOE_RM1 ;
908 _TEMP = WOE_RM1 ;
909 _LPO = _LPO + (          -0.36953980811752 * _TEMP);
910

```

Optimized SAS Code

```

911 *** Effect: WOE_RM2 ;
912 _TEMP = WOE_RM2 ;
913 _LPO = _LPO + ( -0.34433996514399 * _TEMP);
914
915 *** Effect: WOE_RM6 ;
916 _TEMP = WOE_RM6 ;
917 _LPO = _LPO + ( -0.2351286534026 * _TEMP);
918
919 *** Effect: WOE_RM8 ;
920 _TEMP = WOE_RM8 ;
921 _LPO = _LPO + ( -0.24817155145679 * _TEMP);
922
923 *** Effect: WOE_YAS2 ;
924 _TEMP = WOE_YAS2 ;
925 _LPO = _LPO + ( -1.75260650797525 * _TEMP);
926
927 *** Effect: _t_*WOE_GECIKME1 ;
928 _TEMP = _t_ * WOE_GECIKME1 ;
929 _LPO = _LPO + ( -0.07164557184268 * _TEMP);
930
931 *** Effect: _t_*WOE_ORTAK ;
932 _TEMP = _t_ * WOE_ORTAK ;
933 _LPO = _LPO + ( 0.08821660142176 * _TEMP);
934
935 *** Naive Posterior Probabilities;
936 drop _MAXP _IY _PO _P1;
937 drop _LPMAX;
938 _LPMAX= 0;
939 _LPO = -16.6541828432845 + _LPO;
940 if _LPMAX < _LPO then _LPMAX = _LPO;
941 _LPO = exp(_LPO - _LPMAX);
942 _LPMAX = exp(-_LPMAX);
943 _P1 = 1 / (_LPMAX + _LPO);
944 _PO = _LPO * _P1;
945 _P1 = _LPMAX * _P1;
946
947 SURV3DR4:
948
949
950 *** Posterior Probabilities and Predicted Level;
951 label P__g_1 = 'Predicted: _g_1' ;
952 label P__g_0 = 'Predicted: _g_0' ;
953 P__g_1 = _PO;
954 _MAXP = _PO;
955 _IY = 1;
956 P__g_0 = _P1;
957 if (_P1 > _MAXP + 1E-8) then do;
958 _MAXP = _P1;
959 _IY = 2;
960 end;
961 I__g_ = SURV3DRF[_IY];
962 U__g_ = SURV3DRU[_IY];
963
964 *****;
965 ***** end scoring code for regression;
966 *****;
967
968 EM_SUBHZRD1_SURV = P__g_1;
969 EM_SUBHZRD0_SURV =1-(EM_SUBHZRD1_SURV);
970 EM_HZRDFCST=EM_SUBHZRD1_SURV;
971
972 EM_SURVFCST=EM_SURVFCST*(1-EM_HZRDFCST);
973 if _T_=_T0_ then EM_SURVIVAL=EM_SURVFCST;
974 if EM_SURVIVAL >0 then do;
975 if _T_=_t0_fcst then EM_SURVEVENT=(EM_SURVIVAL-EM_SURVFCST)/EM_SURVIVAL;
976 end;
977 else do;
978 if _T_=_t0_fcst then EM_SURVEVENT=(EM_SURVIVAL-EM_SURVFCST)/0.00001;
979 end;
980 _t_+1;

```

```
Optimized SAS Code
980   _t_+1;
981   end;
982   _T_ = _T0_;
983
984   ;
985   end;
986   drop _T0_ t0_fcst
987   _P0 _P1
988   _csb1
989   _csb2
990   _csb3
991   _csb4
992   _csb5
993   BadObs U_g_ I_g_ P_g_ ;
994   *-----*;
995   * TOOL: Score Node;
996   * TYPE: ASSESS;
997   * NODE: Score4;
998   *-----*;
999
```


ÖZGEÇMİŞ

Adı Soyadı : Celil TAŞKIN
Doğum Yeri ve Yılı : Biga, 10/10/1981
Medeni Hali : Evli
Yabancı Dili : İngilizce
E-posta : celiltaskin@gmail.com



Eğitim Durumu

Lise : Çan Sağlık Meslek Lisesi, 1998
Lisans : Boğaziçi Üniversitesi, Fen-Edebiyat Fakültesi, Sosyoloji Bölümü, 2004
Yüksek Lisans : İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 2019

Mesleki Deneyim

Kuveyt Türk Katılım Bankası, Risk Yönetimi Başkanlığı, Kredi Riski ve Modelleme (2014 - devam ediyor)
Kuveyt Türk Katılım Bankası, Kredi Risk İzleme Müdürlüğü, (2012 - 2014)
Kuveyt Türk Katılım Bankası, Kurumsal ve Ticari Krediler Risk Takip Müdürlüğü, (2008 - 2012)

Yayımları

Taşkın, C., Turanlı, M. (2019). TFRS 9 ve Temerrüt Olasılığı Modellemesi. Sosyal Bilimler Araştırma Dergisi, 8 (1), 273-284.