

MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

**TEMERRÜT HALİNDE KAYIP TAHMİNLEMESİ İÇİN
DENETİMLİ VE DENETİMSİZ ÖĞRENME YÖNTEMLERİNİN
BİRLİKTE KULLANILMASI**

YÜKSEK LİSANS TEZİ

Salih SERTBAŞ

İstatistik Anabilim Dalı

İstatistik Programı

Tez Danışmanı: Prof. Dr. Semra ERPOLAT TAŞABAT

06/2021

MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

**TEMERRÜT HALİNDE KAYIP TAHMİNLEMESİ İÇİN
DENETİMLİ VE DENETİMSİZ ÖĞRENME YÖNTEMLERİNİN
BİRLİKTE KULLANILMASI**

YÜKSEK LİSANS TEZİ

Salih SERTBAŞ

İstatistik Anabilim Dalı

İstatistik Programı

Tez Danışmanı: Prof. Dr. Semra ERPOLAT TAŞABAT

06/2021

Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna uygun olarak hazırladığım bu tez çalışmasında;

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel etik 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 değişiklik yapmadığımı,
- ücret karşılığı başka kişilere yazdırmadığımı (dikte etme dışında), uygulamalarımı yaptırmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.



TEMERRÜT HALİNDE KAYIP TAHMİNLEMESİ İÇİN DENETİMLİ VE DENETİMSİZ ÖĞRENME YÖNTEMLERİNİN BİRLİKTE KULLANILMASI

ÖZET

Küresel Ekonomik krizler Bankacılık sektöründe mevcut ve ileriye dönük kredi riskinin doğru ölçülmesi ihtiyacını beraberinde getirmiştir. Bu çerçevede uzlaşılın Basel kararları ile kredi riskine ilişkin Temerrüt Olasılığı (PD), Temerrüt Tutarı (EAD) ve Temerrüt Halinde Kayıp (LGD) gibi risk ölçüm modelleri geliştirilmiştir. Bu modeller sermaye yeterlilik rasyosunun yanı sıra TFRS9 raporlama standartları kapsamında sorunlu kredi karşılık hesaplamalarını da etkilemiştir. Sermaye yeterlilik rasyosu Banka sermayesinin risk ağırlıklı varlıklara oranıdır. Risk ağırlıklı varlıkla ile banka yükümlülüklerinin kayıp riskine esas tutar bulunmaya çalışılır. Bu tutarın hesaplamasında en önemli girdilerden biri temerrüt halinde kayıp oranının (LGD) tahmin edilmesidir. SYR hesaplamaları ile sermaye yeterliliği belirlendiğinden kullanım amacına uygun olarak kısa vadeli konjektörel hareketlerden etkilenmemesi gerekir bu sebep en az bir ekonomik döngü içerisinde yer alan (en az 7 yıl) panel veriler kullanılmalıdır. TFRS9 raporlamasında ise farklı olarak anlık risklerin hesaplanması esastır. Bu nedenle kullanılan panel veri zaman derinliğinden (en fazla 4 yıl) çok yatay genişliği kapsayıcılığı önemlidir.

Kredi riski hesaplamasında kullanılan diğer iki bileşen olan PD ve EAD modelleri için oldukça yaygın bir uygulama ve literatür oluşmasına rağmen LGD modelleri için hem veri hazırlama hem de modelleme tekniği zorlukları sebebiyle bu durum söz konusu değildir. PD ve EAD modelleri için sıklıkla LOGIT Regresyon veya karar ağacı yöntemleri kullanılırken, LGD modelleri için kısıtlı bir literatürle birlikte uygulamada sadece karar ağacı yöntemleri uygulanmıştır. LGD verisi bir müşterinin borçlarını olağan süreçte ödeyemez hale geldikten sonra bankanın alacaklarının ne kadarını tahsil edebileceğini ve bu tahsilatlarının zaman periyodunda ki hareketlerini tahmin etmeye çalışır Ancak panel veri tipine karar ağacı yöntemlerinin uygulanmasında ki zorluklar sebebiyle veri kayıplarına katlanılarak veri basitleştirmelerine gidilmektedir.

Bu çalışmada, LGD modellerinde yapılan yaklaşımların eksik yönleri gösterilerek yenilikçi yöntemler önerilmektedir.



CO-USE OF SUPERVISED AND UNSUPERVISED LEARNING METHODS TO PREDICT LOSS GIVEN DEFAULT

ABSTRACT

Global Economic crises Have led to the need to accurately measure the current and forward credit risk in the banking sector. Within this framework, risk measurement models such as Probability of Default (PD), Default Amount (EAD) and Loss in Default (LGD) related to credit risk have been developed with the basel decisions agreed. In addition to the capital adequacy ratio, these models also affected problem loan provision calculations under TFRS9 reporting standards. The capital adequacy ratio is the ratio of Bank capital to risk-weighted assets. With risk-weighted assets, it is tried to find an amount based on the loss risk of bank liabilities. One of the most important inputs in the calculation of this amount is the estimation of the loss rate (LGD) in case of default. Since the competence to lay is determined by SYR calculations, it should not be affected by short-term congestive movements in accordance with the purpose of use, this reason should be used panel data contained in at least one economic cycle (at least 7 years). In TFRS9 reporting, it is essential to calculate immediate risks. Therefore, horizontal width containerity is more important than the depth of panel data time used (up to 4 years).

Although there is widespread practice and literature for PD and EAD models, the other two components used in credit risk calculation, this is not the case due to both data preparation and modeling technique difficulties for LGD models. While LOGIT Regression or decision tree methods are often used for PD and EAD models, only decision tree methods have been applied in practice with a limited literature for LGD models. LGD data tries to estimate how much of the bank can collect and the movements of its collections over the time period after a customer is unable to pay their debts in the usual process, but data simplifications are made by relying on data losses due to difficulties in implementing decision tree methods to the panel data type.

In this study, innovative methods are proposed by showing the shortcomings of the approaches made in LGD models.



İÇİNDEKİLER

İÇİNDEKİLER	XI
Tablo İçeriği	XIII
Şekil İçeriği	XIV
TEMERRÜT HALİNDE KAYIP TAHMİNLEMESİ İÇİN DENETİMLİ VE DENETİMSİZ ÖĞRENME YÖNTEMLERİNİN BİRLİKTE KULLANILMASI	XVI
CO-USE OF SUPERVISED AND UNSUPERVISED LEARNING METHODS ... TO PREDICT LOSS GIVEN DEFAULT	XVIII
1. GİRİŞ	1
1.1 Temerrüt Halinde Kayıp Üzerine Yapılan Çalışmalar (Literatür Taraması)	5
1.2 Bankacılık Sektörü Tarihsel Gelişimi	7
2. TEMERRÜT VE TEMERRÜT HALİNDE KAYIP (THK)	11
2.1 Temerrüt	11
2.2 Temerrüt Halinde Kayıp (THK)	12
2.3 Tahsilat Eğrisi	13
2.4 THK Dağılım türleri	15
2.4.1 Beta Dağılımı	15
2.5 THK Tahmin Yaklaşımları (Klasik Yöntemler)	16
2.6 THK Tahmini İçin Önerilen Yöntem	19
3. UYGULAMA	21

3.1	Veri Seti ve Değişkenler	21
3.2	Bankalarda Kullanılan Yaygın Yöntem (Klasik Yöntem İle THK Tahmini) ...	33
3.3	Önerilen Yöntem İle THK Tahmini	37
3.3.1	Kümeleme	38
4.	Modellerin Karşılaştırılması	43
5.	Sonuç.....	45
6.	Kaynaklar	47



Tablo İeriđi

Tablo 1. Gcl Ekonođiye Geiř Programı erevesi	8
Tablo 2. Panel Veri dnřm	19
Tablo 3. Veri Setinden bir kesit (Son 10)	21
Tablo 4. Veri Setinden bir kesit (Son 10)-devamı	22
Tablo 5. Veri Seti Deđiřkenlerinin aıklamaları	22
Tablo 6. Ham Eđitim Veri Seti zet İstatistikleri	27
Tablo 7. Trimlenmiř Geliřtirme Verisi zet İstatistikleri	30
Tablo 8. Trimlenmiř Test Verisi zet İstatistikleri	32
Tablo 9. Kategorik Deđiřkenlerin Geliřtirme Verisi zerindeki Dađılımı	33
Tablo 10. Bađımlı Deđiřken rnek Veri Seti	34
Tablo 11. Silhouette Metodu	39
Tablo 12. Elbow Methodu	39
Tablo 13. Klasik ve Yeni Yaklařım RMSE Sonuları	43

Şekil İçeriği

Şekil 1. Tahsilat Eğrileri	14
Şekil 2. Beta Dağılımı	16
Şekil 3. Tahsilat Eğrileri	18
Şekil 4. CCR_TEM_GRUP_3_C değişkenin trimleme öncesi boxplot görünümü	24
Şekil 5. CCR_TEM_GRUP_3_C trimleme sonrası boxplot görünümü	24
Şekil 6. EAD_ADJ değişkenin trimleme öncesi boxplot görünümü	24
Şekil 7. EAD_ADJ trimleme sonrası boxplot görünümü	25
Şekil 8. CCR_TEM_GRUP_5 değişkenin trimleme öncesi boxplot görünümü	25
Şekil 9. CCR_TEM_GRUP_5 değişkenin trimleme sonrası boxplot görünümü	26
Şekil 10. GRUP1_HC_TO_EAD değişkenin trimleme öncesi boxplot görünümü	26
Şekil 11. GRUP1_HC_TO_EAD değişkeninin trimleme sonrası boxplot görünümü	26
Şekil 12. Inter Quartile Range Gösterimi	29
Şekil 13. Örnek Tahsilat Eğrisi	35
Şekil 14. Klasik Model Karar Ağacı	36
Şekil 15. Klasik Karar Ağacı Tahsilat Oranı Eğrileri	37
Şekil 16. Yeni Model Karar Ağacı	41
Şekil 17. Yeni Yaklaşım Karar Ağacı Tahsilat Oranı Eğrileri	42



TEMERRÜT HALİNDE KAYIP TAHMİNLEMESİ İÇİN DENETİMLİ VE DENETİMSİZ ÖĞRENME YÖNTEMLERİNİN BİRLİKTE KULLANILMASI

ÖZET

Küresel Ekonomik krizler Bankacılık sektöründe mevcut ve ileriye dönük kredi riskinin doğru ölçülmesi ihtiyacını beraberinde getirmiştir. Bu çerçevede uzlaşılın Basel kararları ile kredi riskine ilişkin Temerrüt Olasılığı (PD), Temerrüt Tutarı (EAD) ve Temerrüt Halinde Kayıp (LGD) gibi risk ölçüm modelleri geliştirilmiştir. Bu modeller sermaye yeterlilik rasyosunun yanı sıra TFRS9 raporlama standartları kapsamında sorunlu kredi karşılık hesaplamalarını da etkilemiştir. Sermaye yeterlilik rasyosu Banka sermayesinin risk ağırlıklı varlıklara oranıdır. Risk ağırlıklı varlıkla ile banka yükümlülüklerinin kayıp riskine esas tutar bulunmaya çalışılır. Bu tutarın hesaplamasında en önemli girdilerden biri temerrüt halinde kayıp oranının (LGD) tahmin edilmesidir. SYR hesaplamaları ile sermaye yeterliliği belirlendiğinden kullanım amacına uygun olarak kısa vadeli konjektörel hareketlerden etkilenmemesi gerekir bu sebep en az bir ekonomik döngü içerisinde yer alan (en az 7 yıl) panel veriler kullanılmalıdır. TFRS9 raporlamasında ise farklı olarak anlık risklerin hesaplanması esastır. Bu nedenle kullanılan panel veri zaman derinliğinden (en fazla 4 yıl) çok yatay genişliği kapsayıcılığı önemlidir.

Kredi riski hesaplamasında kullanılan diğer iki bileşen olan PD ve EAD modelleri için oldukça yaygın bir uygulama ve literatür oluşmasına rağmen LGD modelleri için hem veri hazırlama hem de modelleme tekniği zorlukları sebebiyle bu durum söz konusu değildir. PD ve EAD modelleri için sıklıkla LOGIT Regresyon veya karar ağacı yöntemleri kullanılırken, LGD modelleri için kısıtlı bir literatürle birlikte uygulamada sadece karar ağacı yöntemleri uygulanmıştır. LGD verisi bir müşterinin borçlarını olağan süreçte ödeyemez hale geldikten sonra bankanın alacaklarının ne kadarını tahsil edebileceğini ve bu tahsilatlarının zaman periyodunda ki hareketlerini tahmin etmeye çalışır Ancak panel veri tipine karar ağacı yöntemlerinin uygulanmasında ki zorluklar sebebiyle veri kayıplarına katlanılarak veri basitleştirmelerine gidilmektedir.

Bu çalışmada, LGD modellerinde yapılan yaklaşımların eksik yönleri gösterilerek yenilikçi yöntemler önerilmektedir.



CO-USE OF SUPERVISED AND UNSUPERVISED LEARNING METHODS TO PREDICT LOSS GIVEN DEFAULT

ABSTRACT

Global Economic crises Have led to the need to accurately measure the current and forward credit risk in the banking sector. Within this framework, risk measurement models such as Probability of Default (PD), Default Amount (EAD) and Loss in Default (LGD) related to credit risk have been developed with the basel decisions agreed. In addition to the capital adequacy ratio, these models also affected problem loan provision calculations under TFRS9 reporting standards. The capital adequacy ratio is the ratio of Bank capital to risk-weighted assets. With risk-weighted assets, it is tried to find an amount based on the loss risk of bank liabilities. One of the most important inputs in the calculation of this amount is the estimation of the loss rate (LGD) in case of default. Since the competence to lay is determined by SYR calculations, it should not be affected by short-term congestive movements in accordance with the purpose of use, this reason should be used panel data contained in at least one economic cycle (at least 7 years). In TFRS9 reporting, it is essential to calculate immediate risks. Therefore, horizontal width containerity is more important than the depth of panel data time used (up to 4 years).

Although there is widespread practice and literature for PD and EAD models, the other two components used in credit risk calculation, this is not the case due to both data preparation and modeling technique difficulties for LGD models. While LOGIT Regression or decision tree methods are often used for PD and EAD models, only decision tree methods have been applied in practice with a limited literature for LGD models. LGD data tries to estimate how much of the bank can collect and the movements of its collections over the time period after a customer is unable to pay their debts in the usual process, but data simplifications are made by relying on data losses due to difficulties in implementing decision tree methods to the panel data type.

In this study, innovative methods are proposed by showing the shortcomings of the approaches made in LGD models.



1. GİRİŞ

Türkiye’de ve dünyada yaşanan yerel ve küresel ekonomik krizler Bankacılık alanında denetimlerin ve denetleyici kuralların sıkılaşmasına yol açmıştır.

Bankacılık sektörünün sık kriz üretmesinin en önemli sebeplerinden biri diğer sektörlerden daha yüksek bir finansal kaldıraç kullanmasıdır. Daralan kredi kanalları aracılığıyla, yaşanan krizler reel kesimde de etkisini göstermiştir. Bu tür olumsuzlukları önleyebilmek, bankacılık denetim, gözetim ve düzenleme kalitesini iyileştirici çalışmalarda bulunmak için uluslararası bankacılık otoriteleri 1974 yılında küresel bir düzenlemenin gerekliliğine karar verilip İsviçre’nin Basel şehrinde bankacılık komitesi toplanarak küresel bankacılık düzenlemelerine bir standart getirmişlerdir. Bankacılık şekil değiştirip küreselleştikçe düzenlemeler de değişmiştir. Özellikle 2009 küresel krizi sonrasında düzenlemelerin önemi daha da artmış ve ek düzenlemeler getirilmiştir.

Basel I, Sermaye Yeterlilik Uzlaşısı adıyla uluslararası piyasalarda işlem yapan tüm bankaların sermaye yeterliliğine ilişkin bir standarttır. Bu standartta ilk aşamada sadece kredi riski için bir sermaye gerekliliği belirlenmiş, izleyen dönemlerde piyasa riski ve diğer bankacılık riskleri için de sermaye tutma zorunluluğu getirilmiştir.

Basel Bankacılık Denetleme Komitesi (BCBS), sermaye yeterliliği ölçümüne ilişkin Basel I’in eksikliklerini ve finansal piyasalarda meydana gelen gelişmeleri dikkate alarak yeni sermaye uzlaşısını duyurarak 2004’te yürürlüğe koymuştur. Bankacılık sektöründe “Basel II” olarak bilinen bu taslak 3 ana bloktan oluşmaktadır.¹

¹ <https://www.bis.org/publ/bcbs107.htm>

Blok 1: Kredi, Piyasa ve Operasyonel riskler hakkında düzenleyici minimum sermaye gereksinimlerini detaylandırmaktadır.

Blok 2: Bankaların sermaye yeterliliğini kendi içsel modelleri ile belirleyebilmesini sağlamaktadır.

Blok 3: Risk ve Sermaye bilgilerinin external taraflara raporlanabilmesi amaçlı rehberler içermektedir.

Basel II'nin getirdiği en önemli yeniliklerden biri bütün bankalara uygulanacak tek tip bir sermaye yeterliliği ölçümünden vazgeçilmesidir.

2006'da operasyonel risk de sermaye yeterliliği hesaplamalarına dahil edilmiştir. Kredi riski hesaplamalarında ise risk duyarlılığı artırılmıştır ve risk ağırlıkları güncellenmiştir. Ayrıca, bankalara bünyelerinde yaptıkları kredi riski analizlerini sermaye yeterliliği raporlarında kullanabilme imkânı sağlanmıştır. (*Saltoğlu, B., 2009*)

Basel II düzenlemelerinin eski anlayış ile temel farklılıkları; sermaye yeterliliği hesaplamasına dahil edilen riskleri çeşitlendirmesi; yani kredi, piyasa ve operasyonel risk yönetimini bir arada gerçekleştiren entegre risk yönetim sistemine geçmesi ve iç denetim ve risk kontrol mekanizmalarına verilen ağırlığı artırması şeklinde değerlendirilebilir. (*Kalkan, A.,2007*)

2008 küresel krizi, mevcut risk düzenlemelerindeki eksikliği ortaya çıkarmıştır. Finansal düzenlemeleri değiştirme kararı alınarak 2009'da Basel 2.5 ortaya çıkmıştır. Piyasa riski ölçümünde getirilen yenilikle riske maruz değer (VAR) tahmini hesaplamasında yaşanmış bir kriz dönemini kapsayan stresli bir dönemi içeren veri kullanılarak sermaye yükümlülüğü hesaplamalarına stresli VAR eklenmiştir.

2008 krizinde banka sermayeleri yetersiz kaldığı için Basel III düzenlenmiştir. Yeni düzenlemeyle makroekonomik istikrara yönelik düzenlemelerle nitelikli sermaye oranı artırılmıştır, likidite yaptırımları eklenmiştir. Sermaye yeterliliği oranlarıyla birlikte uygulanacak bir kaldıraç oranı, döngüsel sermaye tamponu, sermaye koruma tamponu, karşı taraf riski eklenmiştir. Basel III ile ekonomik büyüme hızlandığında kredi büyümesini sınırlayacak, ekonomi yavaşladığında kredi hızını arttıracak mekanizma

kurgulanmıştır. Bu yaklaşımla, bankalara güçlü büyüme dönemlerinde daha fazla sermaye tutma zorunluluğu getirilmiştir. (Saltoğlu, B., 2009)

2013 yılında yeniden gözden geçirilerek “Basel III²” olarak bilinen bu revize taslak aynı zamanda “Avrupa Birliği Sermaye Yeterliliği Yönetmeliği’nde³” ki bankacılık sektöründe “Sermaye Gereksinimleri Yönetmeliği IV³” ya da “CRD IV” olarak bilinen yönetmeliğin de temelini oluşturmaktadır. CRD IV 2014 Ocak ayında yürürlüğe girmiştir ve tüm Avrupa bankaları ve yatırım bankaları için uyulması zorunludur.

Ayrıca Avrupa Bankacılık Otoritesi (EBA), CRD IV gereksinimlerine ilaveten bir dizi ilave rehberler yayınlanmıştır. Özellikle “EBA/GL/2017/16 rehberi⁴” İDD modellerinin geliştirilmesinde bazı spesifik gereksinimler getirmiştir. Bu gereksinimler tüm Avrupa Bankaları tarafından implemente edilmesi gerekmektedir.

Türkiye’de ise bu gelişmelere paralel olarak BDDK tarafından 23.10.2015 tarihinde 29511 sayılı Resmi Gazete’de yayımlanan “Bankaların Sermaye Yeterliliğinin Ölçülmesine ve Değerlendirilmesine İlişkin Yönetmelik⁵” kapsamında bankaların maruz kalınan riskler nedeniyle oluşabilecek zararlara karşı konsolide ve konsolide olmayan bazda yeterli özkaynak bulundurmalarının sağlanmasına ilişkin usul ve esaslar düzenlenmektedir.

Ayrıca BDDK tarafından 23.10.2015 tarihinde 29511 sayılı Resmi Gazete’de yayımlanan “Kredi Riskine Esas Tutarın İçsel Derecelendirmeye Dayalı Yaklaşımlar ile Hesaplanmasına İlişkin Tebliğ⁶” ile bankaların İçsel Derecelendirmeye Dayalı

² <https://www.bis.org/bcbs/publ/d307.htm>

³ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02013R0575-20200627>

⁴ https://eba.europa.eu/documents/10180/2192133/Guidelines+on+PD+and+LGD+estimation+%28EBA-GL-2017-16%29_EN.pdf

⁵ <https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=21192&MevzuatTur=7&MevzuatTertip=5>

⁶ <https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=21194&MevzuatTur=9&MevzuatTertip=5>

Yaklaşımlar kullanarak kredi riskine esas tutar hesaplamalarına ilişkin usul ve esasları düzenlenmektedir.

Özellikle 2005 yılından itibaren Basel 2 kuralları ile Bankacılık alanında kullanılmaya başlanan risk modelleri son 15 yılda oldukça yaygınlaşmıştır. Bankalar beklenmedik risklere karşı kendilerini korumak için yeterli sermaye ayırmak zorundadırlar. Basel III taslağında'da belirtildiği gibi finansal kurumların ayırması gereken sermaye miktarını Risk Ağırlıklı Varlıklar (RAV) formülü baz alınarak aşağıdaki komponentlere bağlı olarak hesaplanır:

Temerrüt Olasılığı (TO),

Temerrüt Tutarı (TT),

Temerrüt Halinde Kayıp (THK),

Efektif Vade (EV)

İleri Düzey İçsel Derecelendirmeye Dayalı Yaklaşımla (A-IRB) Sermaye hesaplamak için TO, TT ve THK değerlerini finansal kurumlar kendi belirledikleri içsel modeller ile hesaplamak durumundadırlar. Tüm risk bileşenleri (TO, TT, THK ve EV) bankalarca hesaplanarak tahmin edilmelidir.

Temerrüt, kelime anlamı hukuk literatüründe direnimdir. Kredi borçlusunun ödeme periyodunda gerekli olan geri ödemeleri bankaya yapmaması durumunda ödeme vadesinden itibaren 90 gün geçmesi durumunda temerrüt sayılmaktadır.

Temerrüt olasılığının ölçülmesi (PD Modelleri) genellikle Logit Regresyon modelleri ile çözümlenen ve kredinin 1 yıl içinde geri ödenmeme ihtimalini ya da Basel tanımıyla 90 gün gecikme yaşanma ihtimalini tahmin eden modellerdir. Genel olarak sermaye piyasalarında mali Analiz tekniklerini içeren uzman görüşüne dayalı bir finansal skorlama literatürü olduğundan PD modellerinin anlamlandırılması yorumlanması bankalar için daha kolay olmuştur.

Bir firmanın 1 yıl içinde temerrüt olma ihtimalinin hesaplanmasından sonra riskin ikinci bacağı olan temerrüt durumunda kredi beklenen büyüklüğünün ölçülmesi gerekmektedir. Temerrüt bakiyesinin (EAD) tahmini, firmaların halen kullanabilecekleri yeni kredilerin

ve mevcut kullanmadıkları limitlerinin bir fonksiyonu olarak değişmekte ve genellikle basit doğrusal regresyon veya karar ağacı yöntemleri ile çözümlenen modellerdir.

Son olarak temerrütte düşen kredi bakiyesi için bankalar borçlarının tahsili için gerek teminatların kullanılması gerek yasal süreçler gerek yeniden yapılandırma ile anlaşma gibi birçok yöntem ile alacaklarını tahsil etmeye çalışmaktadırlar. Temerrüde düştükten sonra başlayan ve dosya bir şekilde kapanana kadar devam eden bu süreçte tahsil edilemeyen bölümün EAD'ye oranı temerrüt halinde kayıp (LGD) olarak adlandırılır.

THK, tahsilat kabiliyetini gösteren bir değer olduğu için kredi riski düzenlemelerinde doğru hesaplanması oldukça önem arz etmektedir. Hem veri işleme zorlukları hem de kavramsal zorluklar yüzünden LGD modelleri Bankacılık literatüründe yaygınlaşmamış ve bilimsel açıdan olgunlaşmamıştır.

1.1 Temerrüt Halinde Kayıp Üzerine Yapılan Çalışmalar (Literatür Taraması)

Kredi riski finansal kurumlar ve onların risk yönetimleri için her zaman önemli bir yer teşkil etmektedir. Bu riski ölçmek ve yönetebilmek hem sermaye açısından hem beklenen kredi kaybı ve de doğru kredi kararları için önem arz etmektedir. Kredi riski komponentinden biri olan Temerrüt olasılığı (TO) hakkında yeteri kadar çalışma olmasına rağmen Temerrüt Halinde Kayıp (THK) hakkında o derecede çalışma bulunmamaktadır. Her ne kadar son yıllarda THK üzerine yapılan araştırmalar artıyor olsa da halen istenilen seviyede değildir. THK ya da onun tamamlayıcısı olan tahsilatlar ile ilgili yapılan bazı çalışmalardan bahsedecek olursak

Gürtler ve Hibbeln (2013) THK'yı ileri düzeyde tahmin edebilen kurumların rekabet açısından bir avantaj sağlayacaklarını söylemiştir.

Schuermann (2004) ise yapmış olduğu çalışmayla tahsilatların 5 karakteristik özelliğini şu şekilde özetler:

Tahsilat dağılımı bimodaldır yani ya çok yüksek ya da çok düşüktür ve düşük oranlı tahsilatlar yüksek oranlı tahsilatlardan daha yaygındır.

Teminat ve derecesi önemlidir. Teminatlı kredilerden elde edilen tahsilat teminatsız kredilerden çok daha fazladır. Dermine ve Carvalho (2005) yılında yaptığı çalışmada

teminat türünün önemli olduğunu ipotek, finansal ve fiziksel teminatların tahsilat oranı anlamında ayrıştığını ancak tüm teminat türlerinin tahsilatlar ile pozitif korelasyonu olduğunu belirtmiştir.

Tahsilatlar ekonomik yapıdan etkilenirler. Yani resesyon olan dönemlerdeki tahsilat oranları büyüme dönemlerindeki göre daha azdır. Frye (2000) yapmış olduğu çalışmada takipteki kredilerinden yapılan tahsilatlar resesyon dönemlerinde %20 azalmaktadır.

Sektör etkisi. Borçlunun faaliyet gösterdiği sektör tahsilat oranını etkilemektedir. Çok fazla miktarda maddi duran varlıklara sahip sektörlerin THK'sı daha az maddi duran varlığı olan sektörlerle (ör:hizmet sektörü) göre daha düşük olmaktadır.

Kredi büyüklüğünün tahsilat oranı modellerinde bir etkisinin olmayacağı belirtilmiştir. Temerrüt olasılığı modellerinde önemli olmasına rağmen tahsilat oranı modellerinde güçlü bir etkisi yoktur.Ancak bazı çalışmalar kredi büyüklüğünün tahsilatları ekileyebileceği belirtilmiştir. Dermine ve Carvalho (2005) yılında yaptığı çalışmada bankaların büyük kredileri temerrüte almaları konusunda spillover effect (yayılma etkisi) den dolayı çok isteksiz oldukları ve büyük ölçekli kredilerin tahsilat performanslarının daha kötü olacağını vurgulamışlardır.

Khieu ve ark (2012) yılında yaptıkları çalışmada kredi türünün de THK oranlarında belirleyici olduğunu göstermiştir. Örneğin;teksitli krediler daha düşük THK değerine sahipken rotatif kredilerde daha yüksektir.

Borçlunun ülkesi,coğrafik konumu da kredi risk modellerinde belirleyici faktör olarak kullanılmaktadır. Çünkü iflas sürecinde ülkelere ya da bölgelere göre değişen yasal farklılıklar THK'yı etkileyebilmektedir. Gupton (2005) ve Araten (2004) coğrafik konumu THK modellerinde kullanmışlardır.

Ayrıca Qi and Yang (2009) temerrüte düşmüş konut kredileri üzerinden THK modeli yaptıklarında en önemli değişkenin LTV (Kredi/Teminat) olduğunu belirlemişler ve modelin uyarlanmış R kare değerinde 0.61 olduğunu, LTV değişkensiz ise 0.15'e düştüğünü görmüşler. Leow ve ark (2009) LTV yanında ekstra bazı değişkenler ile modelin ayrıştırıcı gücünün arttığını da tespit etmişlerdir.

Tez kapsamında THK ele alınarak bankalarda günümüzde uygulanan klasik yaklaşım ile önerilen yaklaşım ele alınarak THK'nın hesaplamasına yeni bir bakış açısı getirilmeye çalışılacaktır. Bu sebeplerle tezde ele alınan eleştiriler ve öneriler finans sektöründe uygulamaya katkılar sunacaktır.

Tez toplam dört bölümden oluşmaktadır. Birinci bölümde Bankacılık Sektörü Tarihsel Gelişimi anlatılırken, ikinci bölümde Temerrüt ve Temerrüt Halinde Kayıp açıklanarak uygulama hakkında detaylı bilgi verilmiş, üçüncü bölümde ise sonuca yer verilmiştir.

1.2 Bankacılık Sektörü Tarihsel Gelişimi

Bankacılık sektörünün tarihsel gelişimi incelendiğinde 24 Ocak 1980 Kararları Türk Bankacılığı'nın gelişmesi açısından son derece önemlidir. Kapalı bir ekonomiden dışa açık bir ekonomiye geçilmiş, ihracata dayalı büyüme politikası benimsenmiştir. Bu dönemde faizlerin serbest bırakılarak piyasada belirlenmesine imkân verilmesi, yerli ve yabancı yeni bankaların kurulmasına olanak verilerek rekabetin artırılması, TDHP'nin (Tek Düzen Hesap Planı) uygulanmaya başlanarak tek ve ortak bir muhasebe sistemine geçilmesi, Sermaye Piyasası Kanunu'nun çıkartılması ve İstanbul Menkul Kıymetler Borsası'nın, TCMB bünyesindeki İnterbank (Bankalar arası Para Piyasası) kurulması sağlanmıştır. (Sümer, 2013: 37)

1990 sonrası dönemde bankaların açık pozisyonlarını yükselterek risksiz ve yüksek getiri sağlayan kamu kâğıtlarına yatırım yaptığı ve asıl bankacılık faaliyetlerinden uzaklaştığı bir dönem yaşanmış, ülkede yaşanan siyasi belirsizliğin etkisi de kırılğan bir yapıda olan sektörü daha da olumsuz etkilemiş ve çok sayıda bankanın TMSF'ye devrinin yaşandığı bir döneme girilmiştir (Günel, 2006). Kasım 2000 döneminde yaşanan krizle birlikte faiz oranları ve enflasyon önemli ölçüde yükselmiş, bu dönemde yüksek döviz açık pozisyonu bulunan bankacılık sektörü likidite riski ve faiz oranları artışı sebebiyle büyük kayıplara maruz kalmış, 1999-2003 dönemleri arasında TMSF bünyesine 20 banka devredilmiş, 8 bankanın faaliyetlerine son verilmiş, 11 bankanın ise birleşmesi gerçekleşmiştir (Kaya, 2012).

Kasım 2000 yılında yaşanan kriz ülke ekonomisi üzerinde olduğu kadar Türk Bankacılık Sektörü üzerinde de oldukça olumsuz etkiler yaratmıştır. Ardından tüm piyasaları etkileyen Şubat 2001 krizinde bankacılık sistemi derinden sarsılmıştır.

Türkiye’de bankacılık sektörünün tarihsel gelişimi incelendiğinde sektörü zayıflatan Şubat 2001 kriziyle birlikte faiz oranları ve döviz kurları hızlı bir şekilde yükselmiştir.

Bankacılık sektörü Şubat 2001 krizi sonrasında hem faiz hem de kur riski sonucu önemli kayıplarla karşı karşıya kalmıştır. Başta kamu bankaları olmak üzere, bankacılık sisteminin yeniden yapılandırılması için ihtiyaç duyulan kaynaklar kamu maliyesi üzerine önemli bir yük getirmiştir.

Türkiye bankacılık alanında düzenlemeler yaparak yapıyı güçlendirmeyi amaçlamış, Nisan 2001’de Güçlü Ekonomiye Geçiş Programı uygulamaya alınmıştır.

Bu programın amacı, kur rejiminin terkedilmesi nedeniyle ortaya çıkan güven bunalımı ve istikrarsızlığı süratle ortadan kaldırmak ve eşanlı olarak bu duruma bir daha geri dönmeyecek şekilde kamu yönetiminin ve ekonominin yeniden yapılandırılmasına yönelik altyapıyı oluşturmaktır. (Türkiye’nin Güçlü Ekonomiye Geçiş Programı)

Tablo 1. Güçlü Ekonomiye Geçiş Programı Çerçevesi

Güçlü Geçiş Unsurları	Ekonomiye Programı	Hedefler	Uygulama Araçları	Uygulama Sonuçları
Yapısal Yenilemeler		Merkez Bankası'nın özerkliği ve piyasa ekonomisinin düzenlenmesi	Merkez Bankası Kanunu, tekellere ve endüstrilerine yönelik Düzenleyici Yasaları	Telekominikasyon, Tütün, Şeker Üst Kurullarının Oluşturulması
Bankacılık Reformları	Sistemi	Sağlam bankacılık sisteminin rasyolarının oluşturulması	Bankacılık Kanunu	BDDK'nın Kurulması, TMSF'ye işlerlik kazandırılması

Faiz ve Kambiyo Politikaları	Yatırımcılar için Orta Vadeli perspektif sağlanması	Kur Sistemleri	Dalgalı Kur Sistemine Geçilmesi, Yönetimli Dalgalanma, Örtülü kur çapasının uygulanması
Ekonomik Stratejileri	Sıkı maliye politikası Anti-enflasyonist yaklaşım, istikrarlı ekonomik büyüme	Yüksek oranlı (%6) Faiz Dışı Fazla, Cari Açık/Milli Gelir oranının kontrolü ve finansmanı	Düşük kur politikasıyla oluşan cari açığın dış borçlanma rakamlarına yansımaları

Kaynak TCMB

2001 sonrası dönemde ithalat artışının ihracatta gerçekleşen artıştan daha yüksek seyretmesi, dış ticaret açığını da beraberinde getirmiştir. 2001-2006 dönemi açısından dış ticaret açığı yüzde 324 oranında artış göstermiş, ihracatın ithalatı karşımıla oranı 2001’de yüzde 75,7 iken dönem sonunda bu oran yüzde 61,8’e gerilemiştir. Dolayısıyla dış ticaret açığının ihracattaki yüksek performansa rağmen ithalattaki artışlardan kaynaklanmış olması muhtemeldir –ki bunun sebebi daha çok iç ekonomik faktörlerden kaynaklanmış ve özellikle ham petrol –doğalgaz gibi madencilik sektörüne ilişkin ithalat gelişmelerinden kaynaklanmıştır. (Ay, 2007: 182)

Dünyanın yaşadığı en önemli finansal krizlerden biri olan Amerika Birleşik Devleti merkezli 2008 finans krizi Lehman Brothers’ın batmasıyla tetiklenmiştir. Lehman Brothers’ın batışında düşük özkaynakla çok büyük riskler alındığı görülmektedir. Küresel boyutta tüm dünyayı etkisini altına alan bu krizde, birçok banka iflas etmiş veya iflasların önlenmesi için bankalar kamusal sermaye altına alınmıştır.

Örneğin, Lehman Brothers’ın batmadan önceki sermaye yeterliliğinin Basel II standartları ile uyumlu olması düzenlemenin yeterliliği konusunda kuşku uyandırmıştır. Bu ve benzer olumsuz gelişmeler risk yönetimi sistemlerine dair düzenleme ve Modelleme eksikliklerinin masaya yatırılması gereğini doğurmuştur. Bu noktada kamu otoriteleri ve düzenleyici otoriteler 2009 G20 toplantılarında Basel II sürecinde aksayan noktaların tekrar gözden geçirilmesi kararı alarak Basel III çalışmalarını başlatmıştır. (Saltoğlu, B., 2009)

Söz konusu kriz ile ulusal ve uluslararası düzenlemelerin yetersizliği ve kriz dönemlerindeki başarısızlığı gözlemlenmiş ve mevcut düzenlemelerin gözden geçirilerek güçlendirilmesine yönelik çalışmalar hızlanmıştır. Bu çerçevede, küresel finans krizi döneminde yoğun eleştirilere konu olan muhasebe standartlarına yönelik olarak IASB tarafından çeşitli çalışmalar yapılmış ve yapılan çalışmalar neticesinde IFRS 9 Standardı yayımlanmıştır. IFRS 9 Standardı ile gerçekleşen kredi zarar modelinin terk edilerek beklenen kredi zararı modeline geçilmesi bankacılık sektöründeki bazı temel uygulamaları değiştirmiş ve özellikle karşılıklar açısından bankacılık sektörü yeni bir boyut kazanmıştır. (Aydın, T., 2020)

01.01.2018 tarihinde TFRS 9'a geçişle birlikte, Kredilerin Sınıflandırılması ve Bunlar için Ayrılacak Karşılıklara İlişkin Usul ve Esaslar Hakkında Yönetmelik yürürlüğe girmiştir.

Diğer taraftan, eski uygulamada genel karşılıklar TFRS 9 ile birlikte yerini, on iki aylık beklenen kredi zarar karşılığı ile borçlunun kredi riskinde önemli sayılabilecek artış nedeniyle ayrılan ömür boyu beklenen kredi zarar karşılığı tutarlarına; özel karşılık ise yerini, borçlunun temerrüde düşmüş olması dolayısıyla ayrılan ömür boyu beklenen Kredi zarar karşılığı tutarlarına bırakmıştır. TFRS 9 ile birlikte beklenen kredi zarar yaklaşımı çerçevesinde bankalar, geliştirdikleri içsel modeller kapsamında kredilerinin temerrütlerine ilişkin olasılık hesaplamaları yapmakta ve beklenen kredi zararı karşılığı hesaplayarak kredilerine karşılık ayırmaktadırlar. Bankalar tarafından geliştirilen içsel modellerde geleceğe yönelik tahminler ve makro değişkenler kullanılmaktadır. (Aydın, T., 2020)

2. TEMERRÜT VE TEMERRÜT HALİNDE KAYIP (THK)

Bu bölümde Temerrüt ve Temerrüt Halinde Kayıp kavramları üzerinde durulacaktır. THK hesaplama yöntemlerine değinilerek tez kapsamında önerilen yöntem detaylandırılacaktır.

2.1 Temerrüt

TFRS 9 doğrudan bir temerrüt tanımı vermemekle birlikte, kredi risk yönetiminde kullanılan ile tutarlı bir şekilde temerrüt tanımı yapılmasını gerektirmektedir.⁷

Kredi Riskine Esas Tutarın İçsel Derecelendirmeye Dayalı Yaklaşımlar ile Hesaplanmasına İlişkin Tebliğ'de yer verilen temerrüt tanımını dikkate alması beklenmektedir. Bahse konu Tebliğ'in EK-2'sinin (29) numaralı paragrafında temerrüt tanımı:

- a) Teminatlara başvurulmaksızın borçlunun bankaya veya konsolide edilen finansal kuruluş niteliğindeki bağlı ortaklıklarından birine olan borçlarını tamamen ödeyemeyeceğine banka tarafından kanaat getirilmesi,
- b) Borçlunun bankaya ve konsolide edilen finansal kuruluş niteliğindeki bağlı ortaklıklarından birine olan önemli tutardaki yükümlülüklerini ifa etmede 90 günden fazla gecikmesi. (*Resmî Gazete, 2015*)

İşletmeler tarafından temerrüt riskinin belirlenmesine yönelik olarak temerrüt tanımları yapılırken, elde edilebilir bilgiler çerçevesinde nitel göstergeler dikkate alınmaktadır. Söz konusu temerrüt tanımı, tüm finansal varlıkları kapsayacak şekilde, tutarlı şekilde uygulanmalıdır.

⁷ https://www.bddk.org.tr/ContentBddk/dokuman/mevzuat_0085.pdf

Diğer taraftan elde edilen bilginin makul ve desteklenebilir olmaması finansal varlığın geri ödemelerinde 90 günü aşması durumu, söz konusu varlığın temerrüte uğradığına ilişkin aksi ispat edilebilir ön koşul olarak kabul edilmektedir. *Aydın, T., 2020*

Temerrüt oranı tahminleri, içsel derecelendirme sisteminde tanımlı her bir kredi derecesine sahip kredi borçlusunun bir yıllık süre içerisinde temerrüde düşme olasılığını gösteren tahminlerdir.

Banka'nın temerrüt tanımı aşağıdaki gibidir:

✓ Borçlunun borcunun vadesini 90 günden fazla geçirmesi,

✓ Bankanın teminatlara başvurulmadan kredi borcunun ödeme yükümlüğünün yerine getirilmesinin mümkün olamayacağı kanaatine ulaşması.

Bu çerçevede temerrüde düşen firma ve müşterilerin temerrüt başlangıç tarihinden dosyanın kapanma tarihine kadar geçen süreç THK modelleri için veri niteliğindedir.

2.2 Temerrüt Halinde Kayıp (THK)

Temerrüt Halinde Kayıp, temerrüde düşen bir kredinin zarar dönüşmesi beklenen kısmıdır. Başka bir deyişle, temerrüde düşen bir krediden dolayı uğranan net kaybın, kredinin temerrüt anındaki bakiyesine oranıdır.

Temel olarak iki tip THK hesaplama yöntemi bulunmaktadır. Bunlar:

1. Piyasa Verisi Kaynaklı THK (Market THK)
2. Teminatın Banka Tarafından Nakde Çevrilmesi (Workout THK)

Piyasa Verisi Kaynaklı THK (Market THK): Bir ürünün piyasa fiyatıyla temerrüt ettiği tarihteki fiyatı üzerinden hesaplanmaktadır.

Kurumsal tahvil fiyatından veya temerrüt noktasında bulunan hisse fiyatından elde edilen tahsilat yüzdesi 1'den çıkarılarak hesaplanır.

Örnek olarak, kötü kredileri satın alan varlık yönetim şirketlerinin fiyatlaması verilebilir. Bir varlık yönetim şirketi 100 milyon TL'lik kötü Kredi portföyünü bir bankadan 30 milyon TL'ye alıyorsa burada THK %70 olarak belirlenebilir. (Saltoğlu, B. 2020)

Teminatın Banka Tarafından Nakde Çevrilmesi (Workout THK): Bir ürüne/krediye ait tasfiye ve izleme süreçlerindeki nakit akışları, temerrüt tarihindeki tutarlarına iskonto edilerek bulunmaktadır. Bu tez kapsamında da “workout” yaklaşımı kullanılmıştır.

Workout THK değerleri iki farklı yaklaşımla hesaplanabilmektedir;

- a. **Gross THK:** Kredi/ürün bazında, krediye/ürüne gelen tüm tahsilatların ve masrafların tek seferde değerlendirilerek net ekonomik kayıp hesaplanarak belirlenen THK'dır.

THK'nın bu hesaplaması, olası veya fiili zarar tutarını, bir kredinin temerrüde düştüğü andaki toplam risk tutarıyla karşılaştırmaktadır.

- b. **Blanco THK:** Ürünlere/kredilere bağlı teminatların net ekonomik kaybı dikkate alınarak teminat tipi kırılımında hesaplanan THK'dır.

Tez kapsamında ele alınan veri setini oluşturan tahsilat verilerinde, tahsilatın hangi teminatlardan yapıldığı bilgisi mevcut değildir. Bu nedenle Blanco THK hesaplaması bu veri deseni ile mümkün olmadığı için THK hesaplamaları Gross THK yöntemi ile yapılmıştır. (tezin bundan sonraki bölümlerinde THK olarak anılacaktır.) THK hesaplamada kullanılan temel formül Eşitlik 1'de gösterildiği gibidir:

$$THK = 1 - \frac{\text{Toplam Tahsilat(indirgenmiş)} - \text{Toplam Masraf}}{\text{Temerrüt Tutarı}} \quad (1)$$

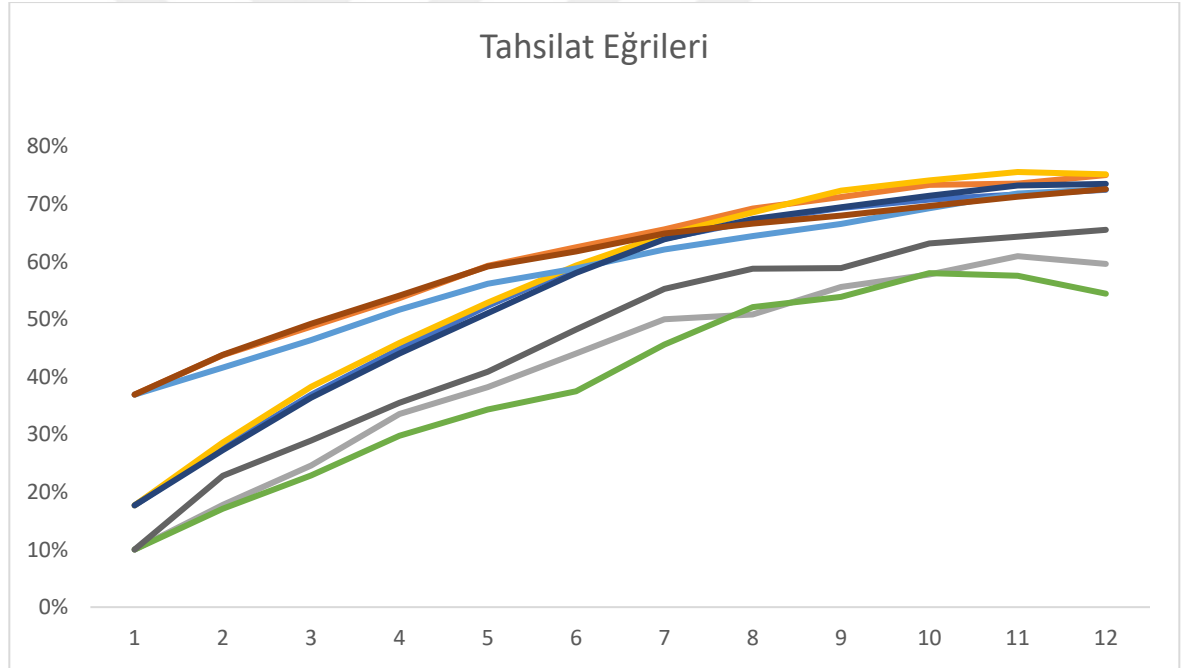
2.3 Tahsilat Eğrisi

Tahsilat eğrisi, takipteki kredilerin temerrütte kalma süreleri bazında gerçekleşen tahsilat oranlarını gösteren eğridir. Başka bir deyişle, takibe giden krediler için yapılan tahsilatların ilgili kredinin faiz oranı ile takip tarihine indirgenerek temerrüt riskine oranının zaman bazında kümülatif olarak gösterimidir.

Eşitlik 1’de gösterildiği gibi 1-THK, toplam net tahsilatın (tahsilat- masraf) temerrüt tutarına oranıdır. Ancak bu tahsilat tek seferde değil zaman içerisinde kesikli olarak yapılan bir süreçtir.

Temerrüde hali hazırda düşmüş ve tahsilatın bir kısmı yapılmış müşterilerin kalan tahsilatlarının tahmin edilebilmesi için THK’nın yanı sıra tahsilat eğrilerinin de tanımlanıp hesaplanması gerekmektedir.

Şekil 1’de farklı tahsilat eğrilerine yer verilmiştir. Her bir müşterinin tahsilat oranları kendi faiz oranıyla takip tarihine indirgenerek ilgili kredinin temerrüt riskine oranının benzer müşterilerle gruplandırılarak tahsilat eğrileri Şekil 1’de gösterilmiştir.



ŞEKİL 1. TAHSİLAT EĞRİLERİ

Temerrüt tutarı sabit olduğundan net tahsilat eğrileri kümülatif devam eden ve sonunda toplam tahsilat oranına diğer bir deyişle 1- THK oranına ulaşılan eğrilerdir. Grafikte yer alan her eğri bir tekil hesaba/müşteriye ait gözlemlerdir.

THK modellerini karmaşık hale getiren ve basit yöntemlerle çözülmesini zorlaştıran nokta sadece toplam tahsilat oranının değil tüm tahsilat sürecinin tahmin edilmesidir. Bu anlamda bir panel veri söz konusudur.

2.4 THK Dağılım türleri

Yapılan ampirik çalışmalar incelendiğinde sıklıkla THK değerleri için weibull, logistic ve beta dağılımları karşılaştırılmıştır. Genel kanı ekonomik beklentilere de uygun olarak iki tepeli Beta dağılımlarının THK verisini daha iyi temsil ettiği yönündedir. Regresyon bazlı modellerde beta regresyon kullanılmaktadır. Ancak Tez kapsamında karar ağacı yöntemlerinden faydalanacağımız için dağılımın kısıtlarını bağlı kalmayacağız.

2.4.1 Beta Dağılımı

THK verisi 0 ile 1 arasında değerler aldığı için bu tür veriler için Beta dağılımı ideal bir dağılım özelliği sunmaktadır. Beta olasılık dağılımı Eşitlik 2’de yer alan formüle sahiptir:

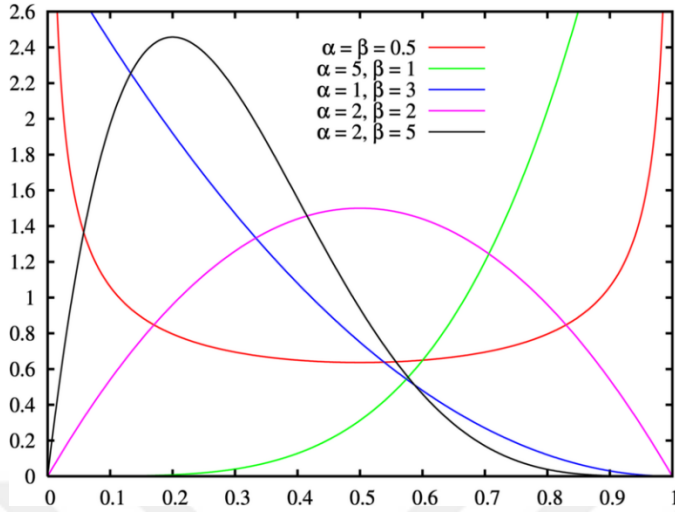
$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \quad (2)$$

Eşitlik 2’de yer alan sembollerin açıklamaları aşağıdaki gibidir:

α : rassal değişkenin 1. Momenti (beklenen değer)

β : rassal değişkenin ölçek parametresi (varyans)

$B(\alpha, \beta)$: hesaplanan olasılık değerinin 0 ile 1 arasında olmasını sağlayan normalleştirme sabiti



ŞEKİL 2. BETA DAĞILIMI

Şekil 2’de farklı α , β parametre değerleri için Beta dağılımının grafiksel gösterimine yer verilmiştir. $\alpha=\beta=0.5$ olarak çizilen U şeklindeki dağılım sıklıkla görülen dağılım şekillerindedir.

2.5 THK Tahmin Yaklaşımları (Klasik Yöntemler)

THK modelleri panel veri olmasına rağmen panel veri regresyonu THK modellemesi için ne akademik çalışmalarda ne de banka uygulamalarında kullanılmamaktadır. Uygulamalarda tahsilat düşünülmeden kümülatif toplam sonuç THK değerleri kesit verisi olarak ele alınarak THK tahmin edilmektedir. Tahmin edilen sonuçlar belirli gruplara ayrılarak zaman boyutunda sonuç tahsilat değeri oluşturan eğrilerin ortalaması beklenen THK değeri olarak kullanılmaktadır.

Akademik çalışmalarda sıklıkla kullanılan regresyon yönteminde veri dağılımlarının olağan En Küçük Kareler ya da Logit regresyon yöntemlerinin varsayımlarını sağlamadaki zorluklar ortaya çıkmaktadır. Genellikle THK değerlerinin beklenen dağılımı bimodal Beta dağılım formundadır. Kuyruk noktalarında aşırı yığılma olduğundan LOGIT gibi binomial regresyon modellerinde sıkça karşılaşılan determinasyon katsayısının düşük ölçülmesi sorununu doğurmaktadır. Ayrıca bağımlı değişkenin normal

dağılımdan oldukça uzak olması hata terimlerinin de normal dağılmaması sorununa yol açmaktadır.

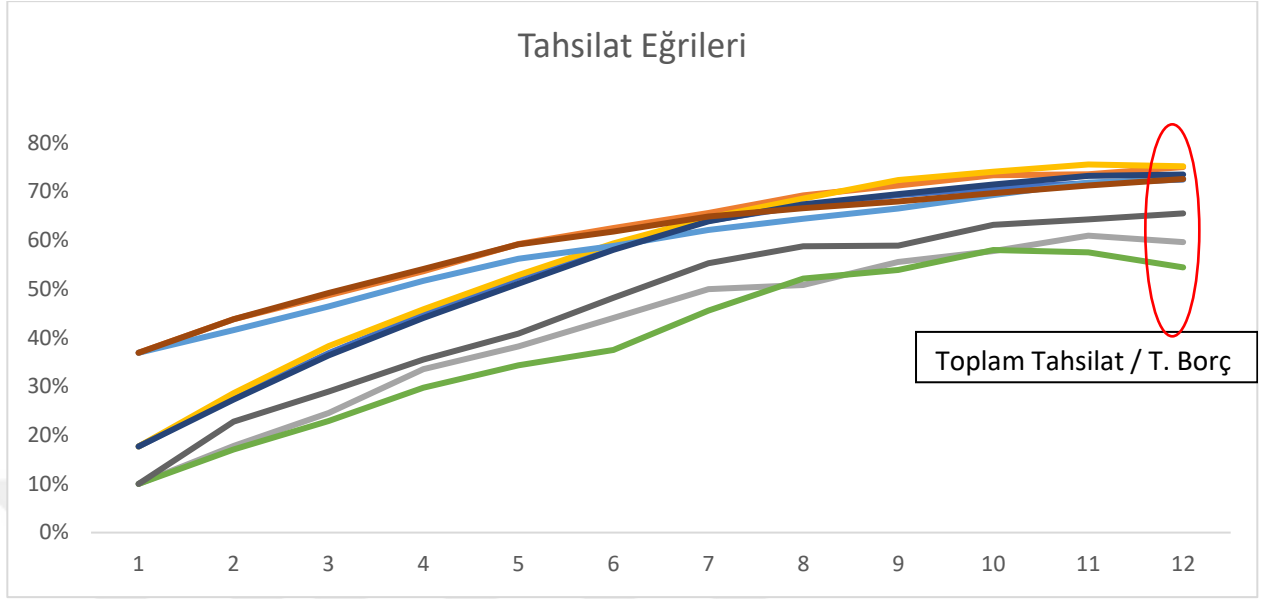
Bu problemlerin çözümü için çeşitli yollar önerilmektedir. Bunların en önemlileri ;

1. Verilerin dönüştürülerek normal dağılıma uygun hale getirilir.
2. İki aşamalı regresyon yönteminde birinci aşamada veriler 0, 1 ve ara değerler olmak üzere 3 kategoriye ayrılarak multinominal logit regresyon yardımı ile sınıflanır. İkinci aşamada ise 0 ve 1 sınıfı haricindeki veriler için klasik doğrusal regresyon yöntemi ile bir tahmin yapılır.
3. Karar ağacı yöntemlerinden faydalanılarak doğrusal olmayan ve normal dağılmayan verilerin regresyonda oluşturduğu dezavantajlar ortadan kaldırılır.

Her halükârda bu yöntemlerin tamamı hedef değişken olarak kümülatif toplam tahsilat oranı üzerinden modellendiği için kümülatif toplam THK rakamının oluşum evrimini dikkate almamaktadır.

Bankalarda uygulamada kullanım kolaylığı sebebiyle karar ağacı yöntemi çok sık kullanılmaktadır. Özellikle Türkiye’de bankalarda çok yaygın olarak SAS yazılımı tercih edilmektedir. SAS yazılımının karar ağacı metodunda “Classification and Regression Trees” (CART) yöntemi herhangi bir özel değişiklik gerektirmeden standart ayarlarda geldiği için çoğunlukla bu algoritma uygulamalarda tercih edilmektedir. Bu yöntem klasik regresyon ile yapılan sınıflama modellerindeki varsayımların birçoğunu içermemesi ve sunduğu parametrik olmayan yapı sayesinde güçlü bir alternatif olmaktadır. Hem kategorik hem de sürekli değişkenlerin sınıflama ve regresyon problemlerinde kullanılabilen CART algoritması her aşamasında ilgili bağımlı değişkeni kendinden daha homojen iki sınıfa ayırarak karar ağacı oluşturur. Ağaçlar gini, twoing, en küçük kareler gibi farklı değişim ölçülerinden yararlanılarak oluşturulabilmektedir. CART algoritması her değişken tipine uygulanabildiği için normallik, dönüşüm gibi işlemleri gerektirmemektedir.

Şekil 3’de yer alan grafikte klasik yöntemdeki vade sonundaki tahsilat oranına odaklanılmaktadır.



Şekil 3. Tahsilat Eğrileri

3 yaklaşım tarzı da tahsilat eğrisinin gelişimi dikkate almadan kümülatif değerler üzerinden modellenmektedir. Yapılan yaklaşım Regresyon modellemesi üzerinden gösterilmek istenirse;

i zaman boyutunu, j kesit boyutunu göstermek üzere;

$Y_{ij} = f(X_{ij}, B) + e_{ij}$, formunda olması gereken panel veri zaman boyutunda kümülatif değerler alınarak kesit verisine dönüştürülmektedir.

$$\sum_{i=0}^n Y_{ij} = f\left(\sum_{i=0}^n X_{ij}, B\right) + e_j \quad (3)$$

(3) numaralı eşitlikte görüldüğü gibi kümülatif değerler üzerinden tahmin edilen her bir Y_j değeri gruplandırılarak (karak ağacı yöntemi ile) her bir grup için hedef değişkenin zaman boyutunda ki tahminleri (4) numaralı formülde görüldüğü gibi grupların ortalaması alınarak tahmin edilir.

$$E(Y_{ij}) = \frac{\sum_{i=0}^n Y_i}{n} \quad (4)$$

Tez kapsamında THK'yı tahminlemek için önerilen yaklaşım yukarıda sıralanan 3. klasik yaklaşım olan vade sonundaki tahsilat oranının CART karar ağacı yöntemi ile tahmin

edilmesi yaklaşımı ile karşılaştırılacaktır. Bu sayede önerilen yöntemin tahmin gücü gözler önüne serilmeye çalışılacaktır. İlerleyen bölümde anlatılacak olan yeni yaklaşım ile tahsilat oranı davranışı benzer kümeler denetimsiz öğrenme algoritmalarından K-means ile belirlenip, belirlenen kümeler yine CART karar ağacı ile THK tahminlenecektir.

2.6 THK Tahmini İçin Önerilen Yöntem

Tez kapsamında THK'yı tahmin etmek amacıyla bankalarda kullanılan klasik yaklaşımlara alternatif bir yöntem önerilmektedir. Önerilen yöntem benzer tahsilat eğrisine sahip grupların denetimsiz öğrenme algoritmaları ile kümelenecek ikinci aşamada kümelerin karar ağacı yöntemiyle tahmin edilmesidir. Panel veriler de ki zaman boyutu kesit noktaları itibari ile faktörlere dönüştürülecek şekilde transpozu alınmaktadır. Böylece her bir panel kesiti tablo 2'de görüleceği üzere tek satır gözleme dönüşmektedir. Böylece her bir zaman faktörü bir kümeleme faktörü olmak üzere gözlemler gruplandırılmaktadır.

Tablo 2. Panel Veri dönüşümü

Bağımlı Değişken	Bağımlı Değişken	t_1	t_2	t_3	...	t_i
$Y_{1,1}$	Y_1	$Y_{1,1}$	$Y_{1,2}$	$Y_{1,3}$...	$Y_{1,i}$
$Y_{1,2}$	Y_2	$Y_{2,1}$	$Y_{2,2}$	$Y_{2,3}$...	
$Y_{1,3}$	Y_3	$Y_{3,1}$	$Y_{3,2}$	$Y_{3,3}$...	
$Y_{1,4}$	Y_4	$Y_{4,1}$	$Y_{4,2}$	$Y_{4,3}$...	
...		\vdots	\vdots	\vdots	...	
$Y_{j,i}$	Y_j	Y_{j1}	Y_{j2}	Y_{j3}	...	Y_{ji}

Tablo 2 de görüldüğü gibi zaman boyutunda transpozu alınarak dönüştürülen veriler kümeleme algoritmaları yardımı ile kümelere ayrılır. Böylece satır bazında tüm veri kümesini temsil eden kategorik bir faktör oluşturulmuş olur.



3. UYGULAMA

Amaç, Temerrüt Halinde Kayıp modellerinde tahmin edilen beklenen tahsilat tutarlarının zaman eğrisi boyunca daha düşük varyanslarla müşteri bazında tahmin edilmesi için geliştirilen hibrit metodun klasik yöntemle karşılaştırılmasıdır.

Uygulamada kümeleme analizleriyle karar ağacı yöntemleri birlikte kullanılmıştır. İlk aşamada tahsilat eğrileri kümeleme analiziyle gruplandırılıp ikinci aşamada her bir grup karar ağacı yöntemi ile açıklayıcı değişkenler kullanılarak tahmin edilmiştir. Bu haliyle yapılan tahmin değerlerindeki tahsilat eğrileri benzer gruplardan geldiği için tahmin varyansları küçültülmüştür.

3.1 Veri Seti ve Değişkenler

Tez kapsamında bir bankaya ait Temmuz 2020-Ocak 2013 tarihleri arasında takibe düşmüş ve takip süreci kapanmış tüzel müşteri verisi ele alınmıştır.

Toplam 22.689 gözlemden oluşan veri setinden ilk 10 kesit verisine Tablo 2’de, son 10 kesit verisine Tablo 3’te yer verilmiştir.

Tablo 3. Veri Setinden bir kesit (Son 10)

ID	LGD	TKP_BAS_T AR	TMRT_BIT_ TAR	EAD	EAD_ADJ	TEMINAT _FLAG	CCR_TEM_ GRUP_1	CCR_TEM_ GRUP_2	CCR_TEM_ GRUP_3_C	CCR_TEM_ GRUP_3_E	CCR_TEM_ _GRUP_4	CCR_TEM_ GRUP_5	GRUP1_HC _TO_EAD	GRUP3C_H C_TO_EAD	GRUP3E_HC_ TO_EAD
4100	0.98	30Sep2015	18Oct2018	1122032	1967329	0	58.5	63.2	52.5	0				-	-
10526	0.02	08Jan2013	30Sep2013	1310419	2824032	1	54.2	74.4	27.4	1.80805731				1.18	1.06
10528	0.80	28Jun2013	27Jun2016	862144	1815846	1	62	34.1	50.9	0.55941212				0.20	0.35
11993	0.01	17Apr2013	24Apr2013	1267109	2697401	0	28.4	70	55.7	0				-	-
16075	0.57	30Oct2015	27Jun2019	689095	1200193	1	21.5	73.5	15.7	1.89435247				0.41	1.10
17992	1.01	11Mar2016	31May2017	50840	86163	0	55.3	61.8	58	0				-	-
18088	0.44	28Jun2013	30Jun2014	15183	31978	0	44.4	16.6	60.4	0				-	-
18476	0.18	30May2014	17Nov2016	304	588	0	60.3	70.9	86.4	0	0	0		-	-
19704	0.12	27Feb2015	07Mar2019	164	301	0	97.5	60.5	31.6	0	0	0		-	-
21256	1.01	30Jan2015	27Jun2016	15650	29067	0	20.9	39	41.7	0				-	-

Tablo 4. Veri Setinden bir kesit (Son 10)-devamı

GRUP4_HC _TO_EAD	GRUP5_HC _TO_EAD	BLACKLIST _FLAG	NEG. INFO	MEMZUC_R ISK_PAYI_1	FARMER	C_RR_ 3	C_RR_ 6	C_RR_ 9	C_RR_ 12	C_RR_ 15	C_RR_ 18	C_RR_ 21	C_RR_ 24	C_RR_ 27	C_RR_ 30	C_RR_ 33	C_RR_ 36	DURASYO N25	DURASYO N50	DURASYO N75	DURASYO N90	
			0	1	0.02	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	37	37	37	37
			0	0	0.30	0	-	0.01	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	7	7	7	7
			1	1	0.37	0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.23	0.23	0.24	37	37	37	37	
			0	0	0.87	0	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1	1	1	1	
			0	1	0.60	0	-	-	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.39	0.39	31	37	37	37	
			0	1	0.04	0	-	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	37	37	37	37	
			0	0	1.00	1	0.14	0.20	0.21	0.25	0.58	0.58	0.58	0.58	0.58	0.58	0.58	13	13	37	37	
			0	-	-	0	-	0.50	0.53	0.54	0.54	1.30	1.30	1.34	1.34	1.34	1.34	4	7	16	16	
			0	-	-	0	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	2.01	2.01	2.01	1	1	1	1	
			0	0	0.05	0	-	-	-	-	0.02	0.02	0.02	0.02	0.02	0.02	0.02	37	37	37	37	

Veri seti %80 eğitim (train) ve %20 doğrulama (test) oranında ikiye ayrılarak model sonuçları doğrulama veri seti ile test edilmiştir. Böylelikle eğitim seti 18.151 doğrulama seti ise 4.538 gözlemden oluşmuştur.

Tablo 5'te veri setinde yer alan değişkenlerin açıklamalarına yer verilmiştir.

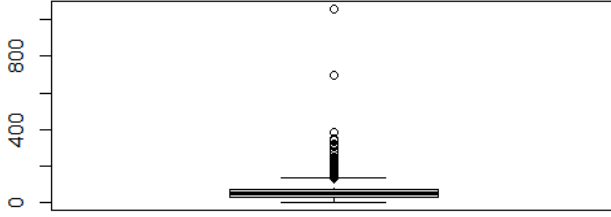
Tablo 5. Veri Seti Değişkenlerinin açıklamaları

Değişken	Açıklama
EAD_ADJ	Temerrüt Tutarı Enflasyon düzeltmesi
EAD	Temerrüt Tutarı
CCR_TEM_GRUP_1	Grup 1 Teminat Tutarı/Toplam Cash Risk
CCR_TEM_GRUP_2	Grup 2 Teminat Tutarı/Toplam Cash Risk
CCR_TEM_GRUP_3_C	Grup 3 Teminat Tutarı/Toplam Cash Risk
CCR_TEM_GRUP_3_E	Grup 3 Expertiz Tutarı/Toplam Cash Risk
CCR_TEM_GRUP_4	Grup 4 Teminat Tutarı/Toplam Cash Risk
CCR_TEM_GRUP_5	Grup 5 Teminat Tutarı/Toplam Cash Risk
TEMINAT_FLAG	Herhangi bir Teminatı var mı yok mu?
GRUP1_HC_TO_EAD	GRUP1_HC/EAD
GRUP3C_HC_EAD	GRUP3C_HC/EAD
GRUP3E_HC_EAD	GRUP3E_HC/EAD
GRUP4_HC_EAD	GRUP4_HC/EAD
GRUP5_HC_EAD	GRUP5_HC/EAD

BLACKLIST_FLAG	Sakıncalılık sorgusu var mı? Varsa 1, yoksa 0
NEG_INFO	Memzuç- Memzuçta olumsuz bir risk bildirimini var mı?
MEMZUC_RISK_PAYI_1	Müşterinin Bankadaki Cash riski/Memzuç toplam Cash riski
FARMER	Çiftçi mi değil mi?
TEMTOPLAM	Teminatların toplam tutarı(TL)
FIRMAAKTIFTOPLAM	Temerrüt anındaki tutar(TL)
C_RR_3	Takip sonrası 3 ay içerisindeki tahsilat oranı
C_RR_6	Takip sonrası 6 ay içerisindeki tahsilat oranı
C_RR_9	Takip sonrası 9 ay içerisindeki tahsilat oranı
C_RR_12	Takip sonrası 12 ay içerisindeki tahsilat oranı
C_RR_15	Takip sonrası 15 ay içerisindeki tahsilat oranı
C_RR_18	Takip sonrası 18 ay içerisindeki tahsilat oranı
C_RR_21	Takip sonrası 21 ay içerisindeki tahsilat oranı
C_RR_24	Takip sonrası 24 ay içerisindeki tahsilat oranı
C_RR_27	Takip sonrası 27 ay içerisindeki tahsilat oranı
C_RR_30	Takip sonrası 30 ay içerisindeki tahsilat oranı
C_RR_33	Takip sonrası 33 ay içerisindeki tahsilat oranı
C_RR_36	Takip sonrası 36 ay içerisindeki tahsilat oranı

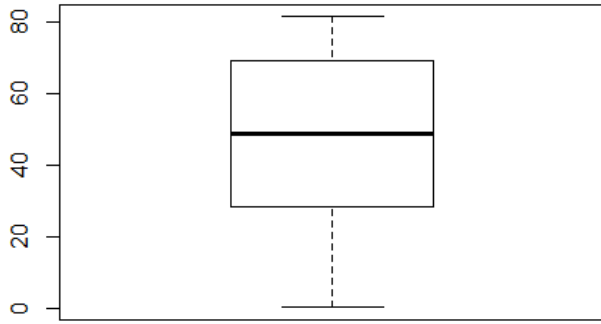
Bir veri kümesinden uç değerleri veya aykırı değerleri kaldırma veya hariç tutma işlemine kırma (trimleme) denmektedir. Aykırı değerler, bir veri kümesindeki dağılımları bozan uç değerlerdir. Uç değerlere sahip değişkenlerde veri setinden ödün vermeden aykırı değerleri ortadan kaldırmak modelleme öncesi başvuru aşamalarından biridir.

CCR_TEM_GRUP_3_C deęişkenin trimleme öncesi boxplot görünümü:



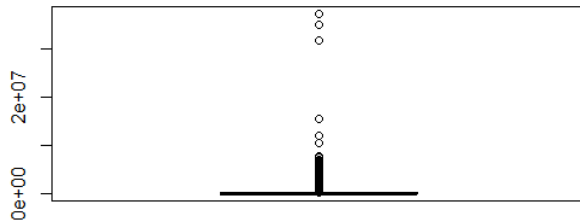
Şekil 4. **CCR_TEM_GRUP_3_C** deęişkenin trimleme öncesi boxplot görünümü görüldüęü gibi veriler pozitif yönlü uç deęer barındırmaktadır.

CCR_TEM_GRUP_3_C trimleme sonrası boxplot görünümü:



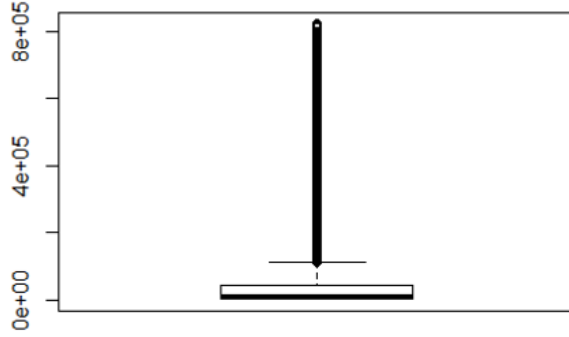
Şekil 5. **CCR_TEM_GRUP_3_C** trimleme sonrası boxplot görünümü görüldüęü gibi deęişken uç deęerlerden arındırılmıştır.

EAD_ADJ deęişkenin trimleme öncesi boxplot görünümü



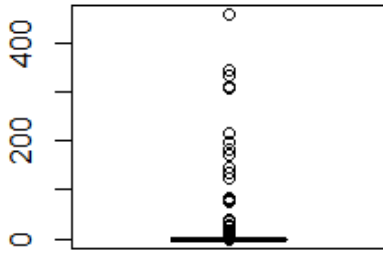
Şekil 6. **EAD_ADJ** deęişkenin trimleme öncesi boxplot görünümü görüldüęü gibi veriler pozitif yönlü uç deęer barındırmaktadır.

EAD_ADJ trimleme sonrası boxplot görünümü



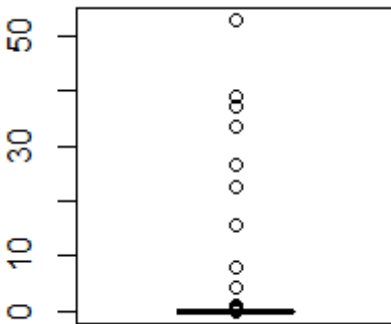
Şekil 7. EAD_ADJ trimleme sonrası boxplot görünümü görüldüğü gibi değişken uç değerlerden arındırılmıştır.

CCR_TEM_GRUP_5 değişkenin trimleme öncesi boxplot görünümü



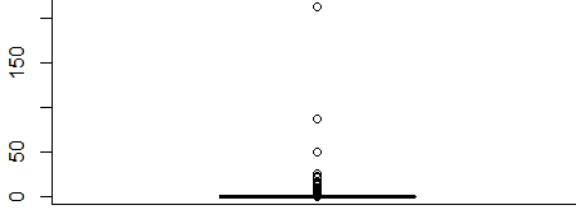
Şekil 8. CCR_TEM_GRUP_5 değişkenin trimleme öncesi boxplot görünümü görüldüğü gibi veriler pozitif yönlü uç değer barındırmaktadır.

CCR_TEM_GRUP_5 değişkenin trimleme sonrası boxplot görünümü



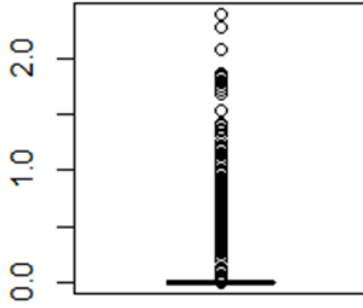
Şekil 9. CCR_TEM_GRUP_5 değişkenin trimleme sonrası boxplot görünümü görüldüğü gibi değişken uç değerlerden arındırılmıştır.

GRUP1_HC_TO_EAD değişkenin trimleme öncesi boxplot görünümü



Şekil 10. GRUP1_HC_TO_EAD değişkenin trimleme öncesi boxplot görünümü görüldüğü gibi veriler pozitif yönlü uç değer barındırmaktadır.

GRUP1_HC_TO_EAD değişkeninin trimleme sonrası boxplot görünümü



Şekil 11. GRUP1_HC_TO_EAD değişkeninin trimleme sonrası boxplot görünümü görüldüğü gibi değişken uç değerlerden arındırılmıştır.

Veri setinin Q1, Q3, Maksimum, Minimum, Ortalama, Medyan, Kayıp değer bilgilerinin yer aldığı özet istatistikleri Tablo 4'te gösterilmiştir.

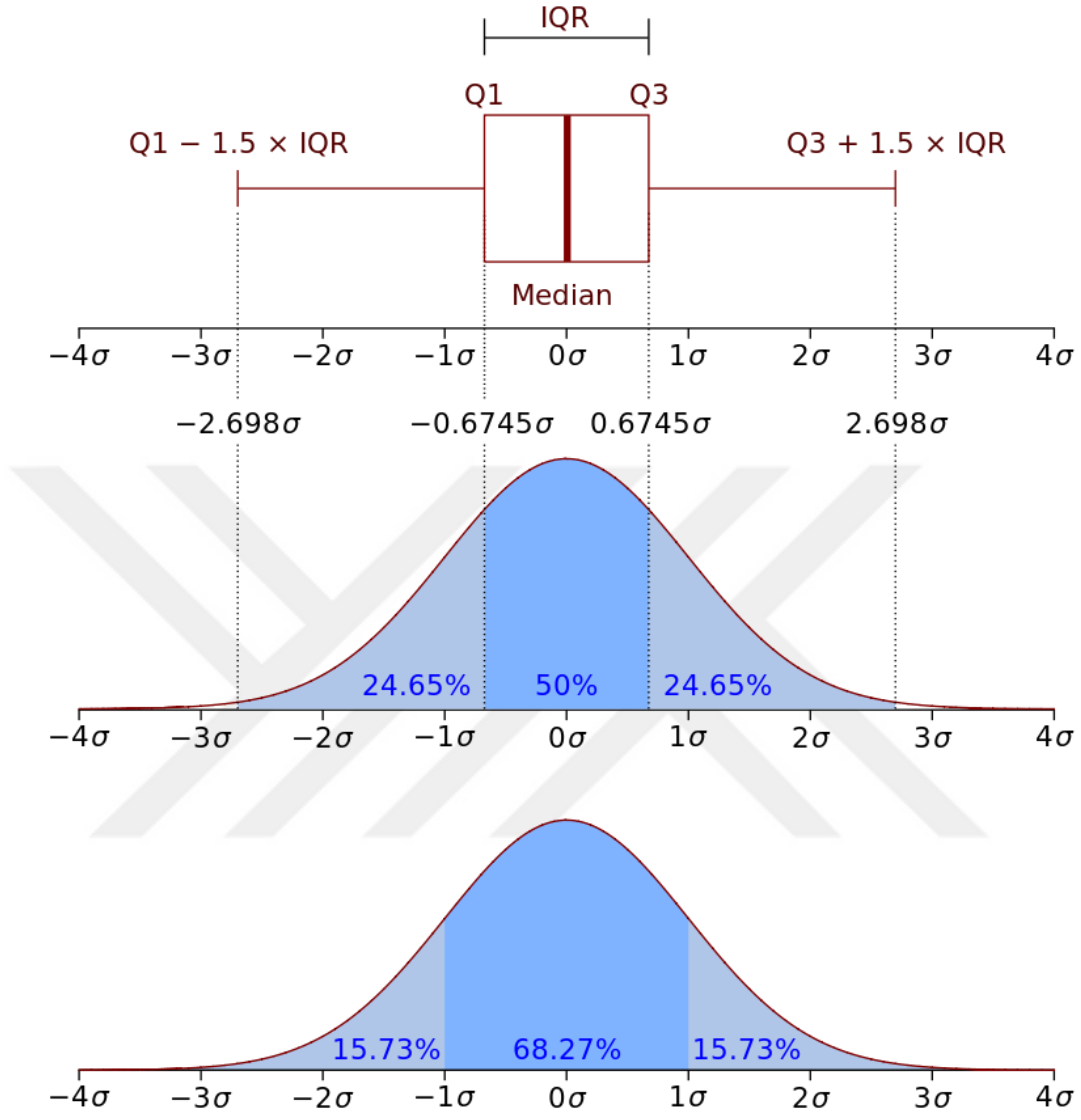
Tablo 6. Ham Eğitim Veri Seti Özet İstatistikleri

Değişken Adı	Q1	Q3	Maksimum	Minimum	Ortalama	Medyan	NA's
C_RR_3	0	0.1778	64.3167	0	0.2004	0	
C_RR_6	0	0.5288	64.3167	0	0.3247	0	
C_RR_9	0	0.83432	102.63726	0	0.44075	0.06439	
C_RR_12	0	0.9375	102.6373	0	0.5516	0.2239	
C_RR_15	0	0.9851	102.6373	0	0.6593	0.4353	
C_RR_18	0.01429	0.99728	102.63726	0	0.74988	0.631	
C_RR_21	0.02128	0.99946	102.63726	0	0.80978	0.71271	
C_RR_24	0.05146	0.99998	102.63726	0	0.85823	0.75896	
C_RR_27	0.08495	1.00837	102.63726	0	0.89949	0.78996	
C_RR_30	0.1027	1.0145	102.6373	0	0.9245	0.8055	
C_RR_33	0.1182	1.0212	102.6373	0	0.9423	0.8198	
C_RR_36	0.1321	1.0263	102.6373	0	0.9555	0.8264	
CCR_TEM_GRUP_1	27.8	69	1075	0.2	49.4	48.7	
CCR_TEM_GRUP_2	28.5	69.1	1046.5	0.1	49.72	49.2	
CCR_TEM_GRUP_3_C	28.2	69.2	1059.2	0.1	49.63	48.8	
CCR_TEM_GRUP_3_E	0	0	21296.296	0	3.688	0	
CCR_TEM_GRUP_4	0	0	5656.109	0	8.357	0	14990
CCR_TEM_GRUP_5	0	0	456.23	0	1.38	0	15454
EAD	343	27201	20755879	100	53700	7483	
EAD_ADJ	572	44331	37197713	106	91567	12120	
GRUP1_HC_TO_EAD	0.508	0.917	212.434	0.001	2.459	0.61	17826
GRUP3C_HC_TO_EAD	0	0	2718.4	0	1.213	0	
GRUP3E_HC_TO_EAD	0	0	3718.544	0	1.762	0	
GRUP4_HC_TO_EAD	0.346	15.172	390.227	0	14.017	0.702	17642
GRUP5_HC_TO_EAD	0.291	59.157	207.688	0.016	36.441	9.949	18106

MEMZUC_RISK_PAYI_1	0.01982	0.59723	1	0	0.3286	0.14066
FIRMAAKTIFTOPLAM	4000000	9000000	11000000	1000000	6037739	6000000
TEMPTOPLAM	21480	48330	59070	5370	32849	32220

Verilerde tespit edilen uç değerlerin etkisini azaltmak için veriler çeyrek açıklığı (interquantile) yöntemi ile kırılmıştır. Çeyrekler açıklığı, bir veri setindeki değişkenlerin ilk çeyrek (Q_1) ve üçüncü çeyrek (Q_3) arasındaki farktır.

Betimsel istatistik ölçütlerinden olan medyan, 1. ve 3. çeyrekler arasında kalan uzaklıklardan faydalanarak uçdeğerlerin tespit edildiği bir yöntemdir. Veriler normal dağıldığında standart sapma ölçütleri ile aynı sonuçlar üretirken özellikle uç değerlerin büyüklüğünden etkilenmemesi sebebiyle normal dağılmayan verilerde çok daha isabetli sonuçlar alınmaktadır. Normal dağılım altında 3. Çeyrek ile 1. çeyrek arasında kalan alanın olasılık karşılığı %50 dir.



Şekil 12. Inter Quartile Range Gösterimi

Çeyrekler arasında kalan alanın 1.5 katı kadar çeyreklerden uzaklaşırsa ortada kalan toplam alan yaklaşık %99.3 ile yaklaşık 3 standart sapma seviyesindedir.

$$IQR = Q_3 - Q_1 \quad (3)$$

$$x_{min} = Q_1 - 1.5 * IQR \quad (4)$$

$$x_{max} = Q_3 + 1.5 * IQR \quad (5)$$

Her bir deęişken ayrı ayrı çeyrekler açıklığı metodu uygulanarak hesaplanan x_{min} deęerinden küçük deęerli kayıtlara x_{min} ve hesaplanan x_{max} deęerinden büyük deęerli kayıtlara ise x_{max} deęişkeni atanmıştır.

Tablo 7. Trimlenmiş Geliştirme Verisi Özet İstatistikleri

Deęişken Adı	Q1	Q3	Maksimum	Minimum	Ortalama	Medyan
C_RR_3	0.000	0.178	1.000	0.000	0.181	0.000
C_RR_6	0.000	0.529	1.000	0.000	0.272	0.000
C_RR_9	0.000	0.834	1.000	0.000	0.347	0.064
C_RR_12	0.000	0.938	1.000	0.000	0.413	0.224
C_RR_15	0.000	0.985	1.000	0.000	0.474	0.435
C_RR_18	0.014	0.997	1.000	0.000	0.525	0.631
C_RR_21	0.021	0.999	1.000	0.000	0.560	0.713
C_RR_24	0.051	1.000	1.000	0.000	0.586	0.759
C_RR_27	0.085	1.000	1.000	0.000	0.605	0.790
C_RR_30	0.103	1.000	1.000	0.000	0.618	0.806
C_RR_33	0.118	1.000	1.000	0.000	0.627	0.820
C_RR_36	0.132	1.000	1.000	0.000	0.633	0.826
CCR_TEM_GRUP_1	27.800	69.000	77.000	0.200	47.290	48.700
CCR_TEM_GRUP_2	28.500	69.100	77.100	0.100	47.630	49.200
CCR_TEM_GRUP_3_C	28.200	69.200	81.600	0.100	48.050	48.800
CCR_TEM_GRUP_3_E	0.000	0.000	4.472	0.000	0.164	0.000
CCR_TEM_GRUP_4	0.000	0.000	38.925	0.000	0.382	0.000
CCR_TEM_GRUP_5	0.000	0.000	52.786	0.000	0.063	0.000
EAD	343.000	27201.000	20755879	100.000	53700.000	7483.000
EAD_ADJ	572.400	44331.500	822749	231.600	60751.800	12120.500
GRUP1_HC_TO_EAD	0.000	0.000	2.393	0.000	0.015	0.000
GRUP3C_HC_TO_EAD	0.000	0.000	4.386	0.000	0.129	0.000

GRUP3E_HC_TO_EAD	0.000	0.000	3.971	0.000	0.118	0.000
GRUP4_HC_TO_EAD	0.000	0.000	39.342	0.000	0.269	0.000



Tablo 8. Trimlenmiş Test Verisi Özet İstatistikleri

Değişken Adı	Q1	Q3	Maksimum	Minimum	Ortalama	Medyan
C_RR_3	0.000	0.175	1.000	0.000	0.180	0.000
C_RR_6	0.000	0.528	1.000	0.000	0.272	0.000
C_RR_9	0.000	0.856	1.000	0.000	0.354	0.069
C_RR_12	0.000	0.949	1.000	0.000	0.414	0.212
C_RR_15	0.000	0.988	1.000	0.000	0.473	0.426
C_RR_18	0.014	0.998	1.000	0.000	0.523	0.608
C_RR_21	0.026	0.999	1.000	0.000	0.556	0.695
C_RR_24	0.050	1.000	1.000	0.000	0.581	0.743
C_RR_27	0.084	1.000	1.000	0.000	0.599	0.773
C_RR_30	0.095	1.000	1.000	0.000	0.609	0.791
C_RR_33	0.112	1.000	1.000	0.000	0.620	0.805
C_RR_36	0.122	1.000	1.000	0.000	0.625	0.815
CCR_TEM_GRUP_1	27.600	68.900	76.900	0.200	47.172	49.150
CCR_TEM_GRUP_2	27.900	69.300	76.900	0.100	47.525	49.100
CCR_TEM_GRUP_3_C	27.600	68.900	81.400	0.200	47.531	48.300
CCR_TEM_GRUP_3_E	0.000	0.000	4.183	0.000	0.159	0.000
CCR_TEM_GRUP_4	0.000	0.000	38.661	0.000	0.487	0.000
CCR_TEM_GRUP_5	0.000	0.000	57.560	0.000	0.105	0.000
EAD	356	26806	8924036	100	52866	7284
EAD_ADJ	592	43361	831657	232	60652	11858
GRUP1_HC_TO_EAD	0.000	0.000	3.201	0.000	0.014	0.000
GRUP3C_HC_TO_EAD	0.000	0.000	3.697	0.000	0.121	0.000
GRUP3E_HC_TO_EAD	0.000	0.000	3.578	0.000	0.113	0.000
GRUP4_HC_TO_EAD	0.000	0.000	39.496	0.000	0.340	0.000

Trimleme öncesinde bazı değişkenler aykırı değerlerden arındırılarak trimleme sonrası değişkenlerle veri modelleme için hazırlanmıştır.

Tablo 9. Kategorik Değişkenlerin Geliştirme Verisi Üzerindeki Dağılımı

Değişken Adı	Müşteri Sayısı		
	0	1	NA
TEMINAT_FLAG	15673	2478	
BLACKLIST_FLAG	14528	971	2652
NEG_INFO	13972	4179	
FARMER	11605	6546	

Tablo 9’da 0 (yok), 1 (var) şeklinde işaretlenen değişkenlerin müşteri sayısı dağılımı görülmektedir.

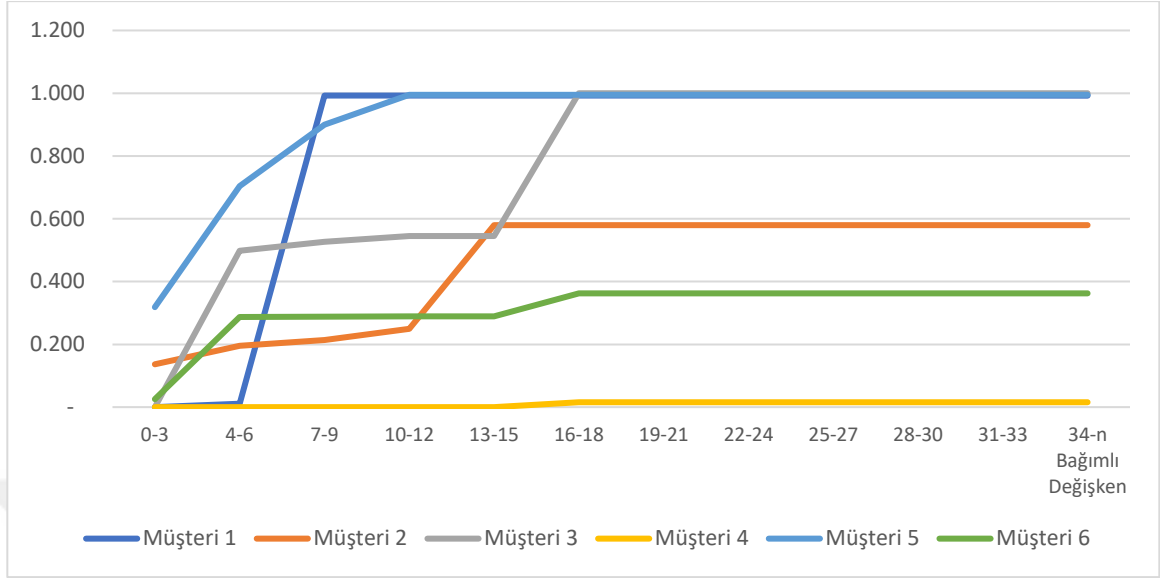
3.2 Bankalarda Kullanılan Yaygın Yöntem (Klasik Yöntem İle THK Tahmini)

Bölüm 2.5. *THK Tahmin Yaklaşımları (Klasik Yöntemler)*’te açıklanan CART algoritması klasik yöntem uygulamalarında sıklıkla karar ağacı yönteminde kullanılarak toplam tahsilat oranı belirli açıklayıcı değişkenler ile ayrıştırılmaktadır. Her bir sonuç sınıfında oluşan tahsilat oranları mevcut değişkenler ile ayrıştırılabilecek en homojen oranları barındırmaktadır. Böylece ağacın birbirinden farklı dallarında oluşan düğümlerindeki firmaların ortalama tahsilat oranı o düğüme düşen firmaların tahmini beklenen tahsilat oranı olmaktadır. Aynı zamanda her bir düğümde yer alan firmaların tahsilat eğrilerinin ortalaması da o düğümdeki firmaların beklenen tahsilat eğrisi olarak hesaplanmaktadır. Ancak tahsilat oranlarının homojen olması tahsilat eğrilerinin de homojen olmasını garanti etmemektedir. Bu dezavantaj verideki panel yapının dikkate alınmamasından kaynaklanmaktadır. Uygulamada kullanılacak veri setinin bağımlı değişkeni olan kümülatif ve toplam tahsilat oranı örneklerine aşağıda yer alan Tablo 10’da yer verilmiştir.

Tablo 10. Bağımlı Değişken Örnek Veri Seti

Müşteri	Zaman											34-36 Bağımlı Değişken
	0-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-30	31-33	
1	-	0.011	0.992	0.992	0.992	0.992	0.992	0.992	0.992	0.992	0.992	0.992
2	0.137	0.195	0.214	0.249	0.580	0.580	0.580	0.580	0.580	0.580	0.580	0.580
3	-	0.498	0.526	0.545	0.545	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	-	-	-	-	-	0.016	0.016	0.016	0.016	0.016	0.016	0.016
6	0.319	0.705	0.900	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
7	0.025	0.287	0.288	0.289	0.289	0.362	0.362	0.362	0.362	0.362	0.362	0.362

Tablo 10’da da görüldüğü üzere THK modellerinde hedef değişken panel verisidir. Ayrıca zaman boyutu itibari ile kümülatif bir veri olduğundan zamana göre otokorelasyonlu bir veridir. Uygulamada panel verisi ile çalışma zorlukları sebebiyle modellerde hedef değişken olarak nihai kümülatif toplam olan “34-n” değişkeni hedef değişken olarak kullanılır. Bu değişken aynı zamanda 1-Kayıp oranı (THK) olarak adlandırılır. Genellikle karar ağacı yöntemi ile tahmin edilen THK oranları için ağacın her bir düğümüne denk gelen müşterilerin diğer aylardaki kayıp oranları da o gözlemlerin ortalamasından hesaplanmaktadır.



Şekil 13. Örnek Tahsilat Eğrisi

Şekil 13'te örnek olarak verilen sınırlı sayıda gözlemde "34-n" yani kümülatif toplam tahsilat oranı hedef değişken olarak seçilerek karar ağacı yöntemi ile düğümler bulunmuştur. Artık bu düğümlerdeki ortalama "34-n" değeri hedef değişkenin beklenen değeri kabul edilir. Aynı düğümlerdeki gözlemlerin kümülatif tahsilat eğrilerinin ortalamasında ortalama tahsilat eğrisinin beklenen değeri kabul edilmektedir.

$$\sum_{i=1}^n LGD_{ay,düğüm}/N \quad (6)$$

Bankalarda kullanılan klasik yöntemi test etmek için kullanılacak karar ağacı R programlama dilinde "Rpart" paketi kullanılarak CART algoritması ile oluşturulmuştur. Algoritma parametrelerine aşağıda yer verilmiştir:

```
rpart(train$C_RR_36 ~. ,train[,c(6:22)], control = rpart.control(minsplit = 100, cp=
0.00005, maxcompete = 5, usesurrogate = 2, maxdepth=4, xval= 10, method ="anova")
```

Yukarıdaki R kodunda belirtilen ifadeler aşağıda açıklanmaktadır:

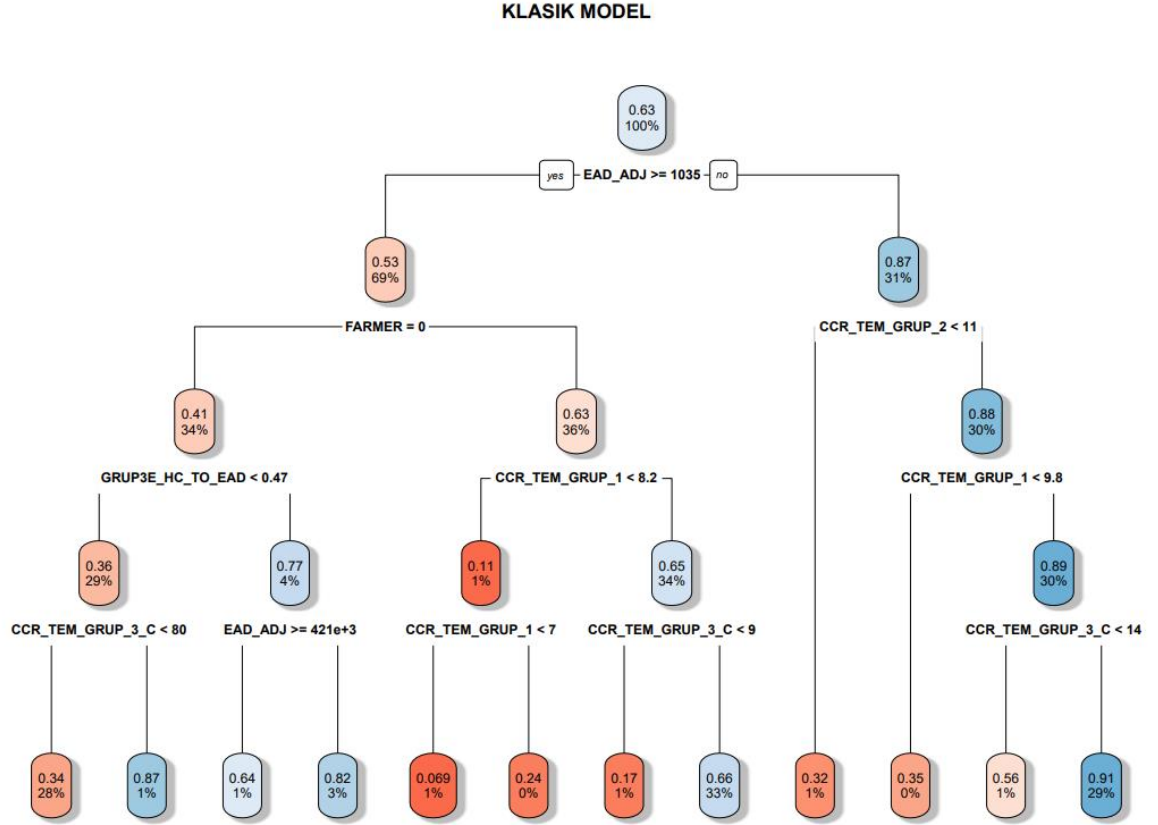
Minsplit: Her düğümden bulunması gereken en düşük gözlem sayısıdır.

cp: Her bir düğüm bölünmesinin sağlanması için başarı ölçütünde ki beklenen en düşük artış oranı

Maxcompete : Çıktıda tutulan rakip bölmelerinin sayısı.

Maxdepth : En fazla oluşabilecek ağaç derinliği

Xval: Çapraz validasyon sayısıdır.

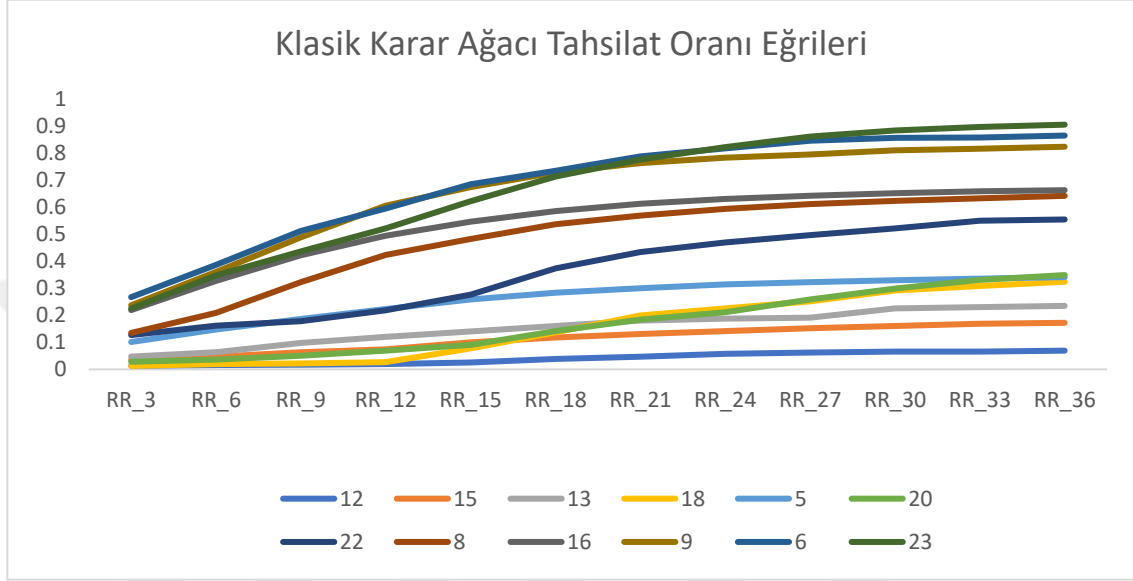


Şekil 14. Klasik Model Karar Ağacı

Ortalama değeri %63 Toplam tahsilat oranı 4 basamakta toplam 12 farklı homejen gruba ayrılmıştır. Her bir grup belirlenen kriterler doğrultusunda bir alt gruba heterojen biçimde bölünemediği için grupların ortalama tahsilat oranları o grupta ki gözlemlerin beklenen en iyi tahsilat oranıdır.

İkinci aşamada her bir gruptaki gözlemin toplam tahsilat oranını oluşturan kümülatif tahsilat eğrilerinin ortalaması alınmaktadır. Bu ortalama tahsilat eğrileri o gruba girecek gözlemlerin tahsilat eğrisini temsilen kullanılır.

Karar ağacının son düğümlerinde yer alan her bir kümenin ortalama tahsilat eğrileri Şekil 15'teki gibidir.



Şekil 15. Klasik Karar Ağacı Tahsilat Oranı Eğrileri

3.3 Önerilen Yöntem İle THK Tahmini

Mevcut yapıda karar ağaçları hedef değişken olan kümülatif toplam tahsilat oranına yani tahsilat eğrisinin son noktasına odaklanmakta eğrinin geri kalanını için ilgili sınıfın ortalamasını almaktadır. Bu sebeple toplam tahsilat oranı aynı olan ancak tahsilat eğrileri farklı olan kitleleri yeterince ayırtıramamaktadır. Genel olarak bankalarda toplam tahsilat oranının tahmini henüz temerrüde düşmemiş firmalarda önemli iken hali hazırda tahsilat sürecinin için de olan temerrüt firmaları için tahsilat eğrisinin tahmini daha önemli olabilmektedir. Bu sebeple hem panel verinin gerektirdiği ileri seviye test ve analizlere girmeden hem de bilgi kaybını en aza indirmek için modelleme öncesinde tahsilat eğrileri için bir kümeleme analizi yapılarak devamında her bir kümenin bir karar ağacı algoritması kullanılarak tahmin edilmesi yöntemi test edilecektir.

Denetimsiz öğrenme modellerinden K-ortalama kümeleme analizi sayesinde benzer kümülatif tahsilat eğrilerine sahip gözlemler kümelendi. İkinci aşamada karar ağacı yöntemi ile (CART) kümeler açıklayıcı değişkenler vasıtası ile sınıflanmaktadır. Böylece

oluşturulan model sadece toplam tahsilat tutarının değil tüm kümülatif tahsilat eğrisi benzer gözlemlerin bulunduğu kümeleri tahmin edebilir hale gelmektedir. 3. Aşamada ise karar ağacında tahmin edilen her bir kümenin ortalama tahsilat oranlarının ve eğrilerinin hesaplanması ile beklenen tahmin değerleri elde edilmektedir.

Bu yöntem ile klasik yöntem nazaran elde edilen en büyük avantaj toplam beklenen tahsilat oranı tahmin edilen sınıftaki gözlemlerin kümülatif tahsilat eğrilerinin de benzer olmasıdır. Elde edilen homojenlik sayede hesaplanan ortalama tahsilat eğrilerinin de varyansı azalmış olmaktadır. Banka tahsilat politikaları açısından ise belirli aksiyon planlarının benzer müşteriler üzerinde uygulanabilmesi sağlanmaktadır.

3.3.1 Kümeleme

K Means (K-Kümeleme) bir veri setinde benzer nicelikler taşıyan gözlemleri kümelemek amacıyla önsel olarak verilen k parametresi adedince küme oluşturmaktadır. Kümeler oluşturulurken dikkate alınan amaç/maliyet fonksiyonu kümeler arası benzerlikleri en düşük yaparken kümeler arası farklılıkları maksimum yapmaktır. Gözlemler THK değerleri tahmin edilen müşterileri, nicelikler ise o müşterilerin zaman boyutundaki tahsilat eğrileridir. Böylece Tablo 1'de verilen bağımlı değişken verileri kümelenebilir.

Silhouette skoru, K-Means gibi kümeleme algoritmaları kullanılarak oluşturulan kümelerin kalitesini, örneklerin birbirine benzer diğer örneklerle ne kadar iyi kümelendiğini değerlendirmek için kullanılır. Silhouette puanı, farklı kümelerin her bir örneği için hesaplanır. Her gözlem/veri noktası için Silhouette puanını hesaplamak için, tüm kümelere ait her gözlem için aşağıdaki mesafelerin bulunması gerekir:

- Aynı kümedeki gözlem ve diğer tüm veri noktaları arasındaki ortalama mesafe. Bu uzaklığa ortalama küme içi uzaklık da denilebilir. Ortalama mesafe bir ile gösterilir.
- Gözlem ile bir sonraki en yakın kümenin diğer tüm veri noktaları arasındaki ortalama mesafe. Bu mesafe aynı zamanda ortalama en yakın küme mesafesi olarak da adlandırılabilir. Ortalama mesafe b ile gösterilir

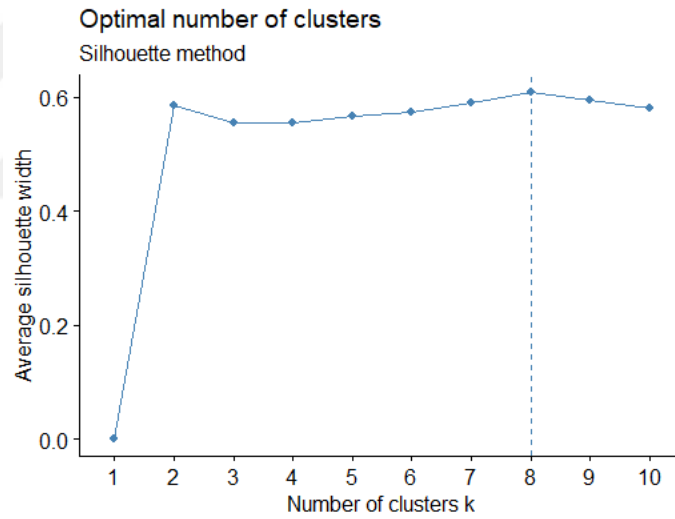
Silhouette puanı, S, her numune için 7 numaralı formül kullanılarak hesaplanmaktadır:

$$(S = \frac{(b - a)}{\max(a, b)}) \quad (7)$$

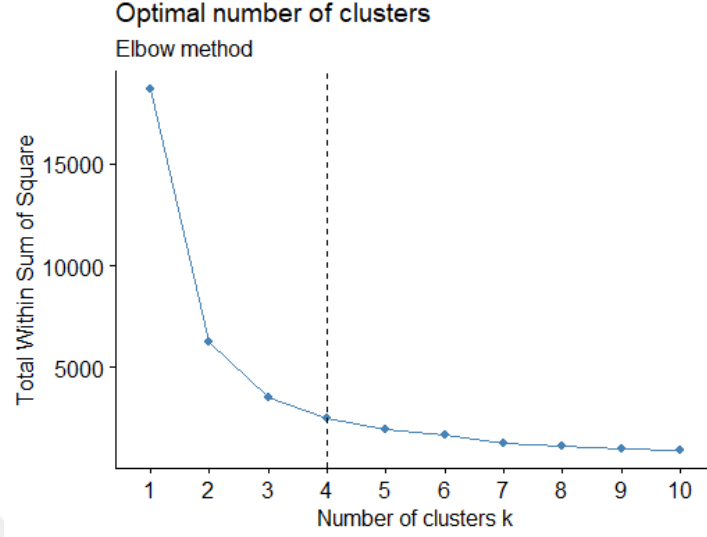
Elbow grafiğinde ise küme içi kareler toplamının küme sayısının bir fonksiyonu cinsinden gösterilmesidir. Böylece her bir ek kümelemenin getireceği katkı sezgisel olarak incelenir ve en uygun küme sayısı belirlenir.

Tablo 11’de gösterilen aylar itibari ile ölçülen her kümülatif tahsilat değeri bir değişken kabul edilerek K-means algoritması ile segmente edilmektedir. Silhouette ve Elbow methodun önermiş olduğu segment sayıları aşağıdaki gibidir.

Tablo 11. Silhouette Metodu

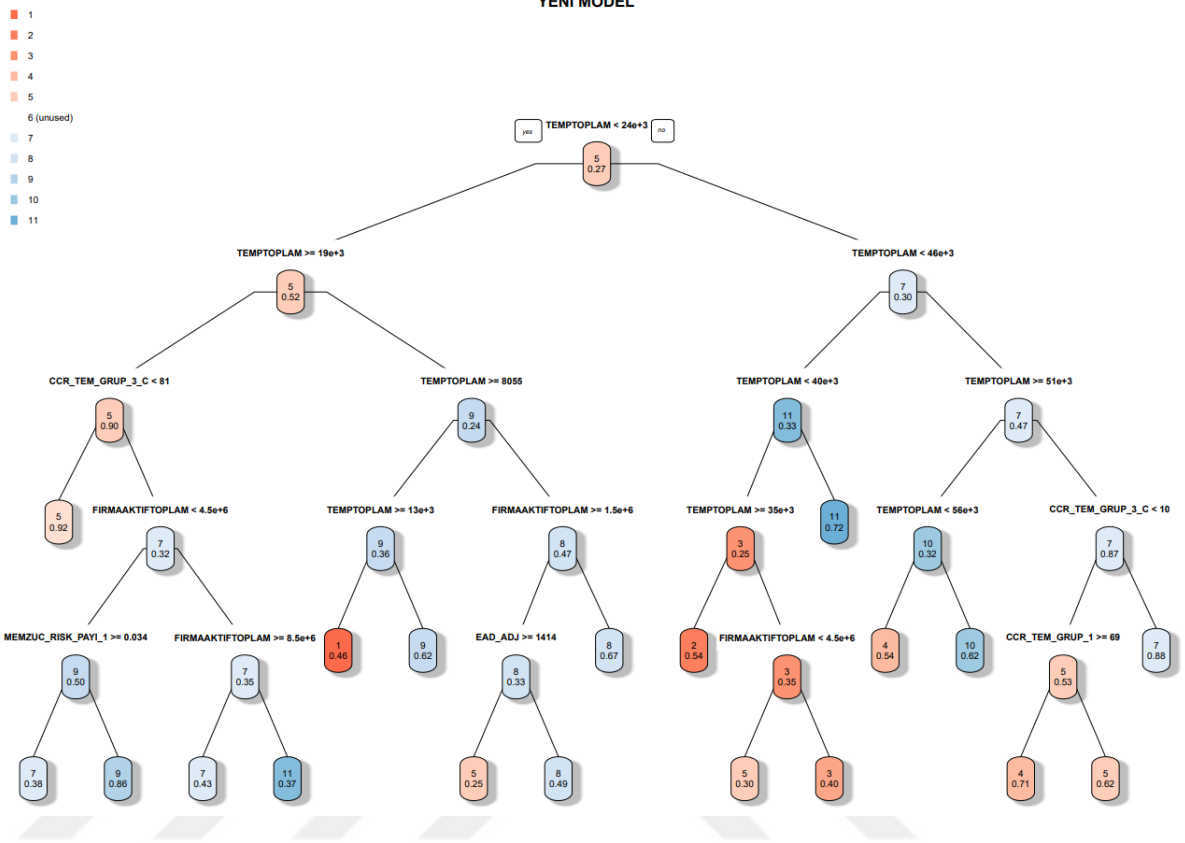


Tablo 12. Elbow Methodu



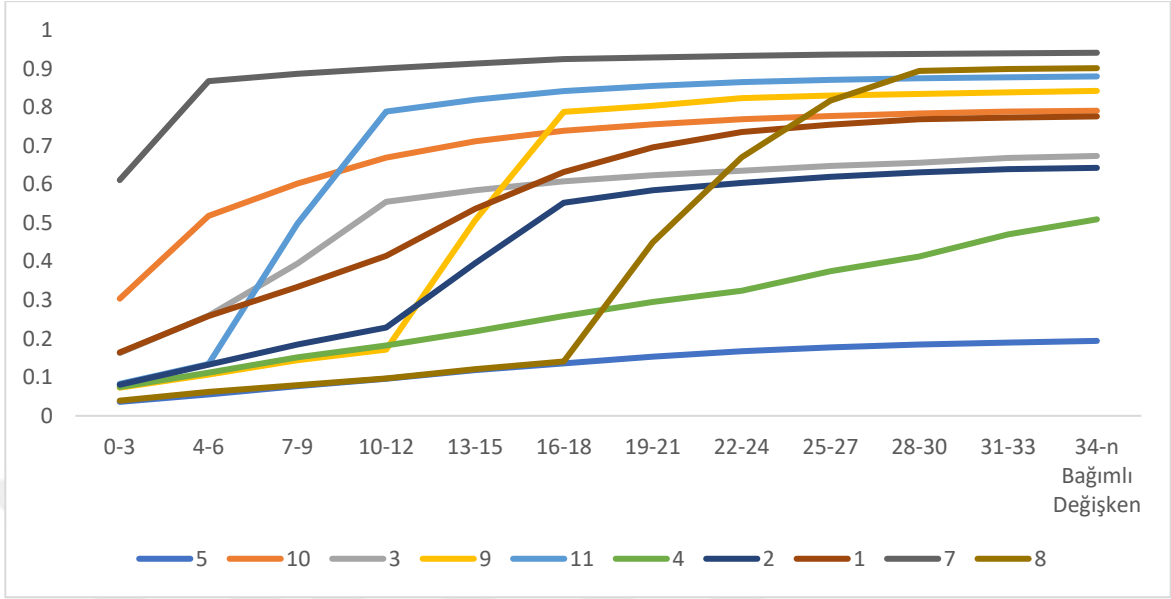
Elbow 4 , Silhouette 8 küme önermektedir. Ancak ayrıştırma katkısı düşük gereksiz kümeler ilgili gözlemlerin açıklayıcı değişkenleri açısından da benzerse karar ağacı modelinde tekrar aynı sınıflarda toplanacağından küme sayısı olabildiğince geniş tutulmuştur böylece bağımlı değişkeni sürekli seriden kategorik veriye dönüştürülmesi ile yaşanan bilgi kaybı en aza indirilecektir. Bu sebeple gözlemler yani tahsilat eğrileri K means yöntemi kullanılarak 11 kümeye ayrılmıştır.

Toplam tahsilat oranı yerine tüm tahsilat noktaları ile yapılan k means kümeleme çalışması sonrasında oluşturulan kümelerin tahmin edildiği karar ağacı aşağıdaki gibidir.



Şekil 16. Yeni Model Karar Ağacı

Karar ağacı yaklaşımında toplam 19 farklı sonuç düğümü oluşturulmuştur. Ancak belirlenen 11 kümeden sadece 1'i karar ağacında hiç bir sınıfata yer almamıştır. 2. Aşamada her bir sınıfın ortalama toplam tahsilat oranı ve ortalama tahsilat eğrileri belirlenmiştir.



Şekil 17. Yeni Yaklaşım Karar Ağacı Tahsilat Oranı Eğrileri

4. Modellerin Karşılaştırılması

Her iki yöntemle elde edilen modeller test veri üzerinde çalıştırılarak test tahmin verileri elde edilmiştir. Elde edilen tahmin verilerinin performansı Kök Ortalama Karekök Hatası (RMSE) kullanılarak karşılaştırılmıştır. RMSE tahmin hatalarının standart sapmasıdır. Düşük RMSE değerine sahip model performansı daha iyi olduğu sonucunu vermektedir. Formül 8’de RMSE hesaplamasına yer verilmiştir.

$$\text{RMSE (Root Mean Squared Error)} = \sqrt{\frac{\sum_{i=1}^N (y - \hat{y})^2}{N}} \quad (8)$$

Formül 8’de yer alan değişken açıklamaları aşağıda açıklanmıştır:

y: Gerçekleşen THK

\hat{y} : Tahmini THK

Tablo 13. Klasik ve Yeni Yaklaşım RMSE Sonuçları

	Klasik Yaklaşım RMSE	Yeni Yöntem RMSE
C_RR_3	32.63%	27.18%
C_RR_6	37.79%	26.36%
C_RR_9	40.09%	30.14%
C_RR_12	40.78%	29.82%
C_RR_15	40.26%	31.06%
C_RR_18	39.04%	29.87%
C_RR_21	37.61%	30.63%

Tüm dönemlerde yeni yöntem ile elde edilen model klasik yaklaşımdan daha düşük RMSE değeri üretmiştir.

C_RR_24	36.36%	30.26%
C_RR_27	35.09%	29.39%
C_RR_30	34.30%	28.81%
C_RR_33	33.50%	28.68%
C_RR_36	33.16%	28.51%

Beklenen Tahsilat (Expected Recovery)

Formul 9’da beklenen tahsilat formülüne yer verilmektedir.

$$\mathbf{Beklenen\ Tahsilat : (1 - LGD) * EAD \quad (9)}$$

Modelleri karşılaştırmak amacı ile THK değeri yerine tahmin edilen C_RR_36 THK değerleri kullanılacaktır.

Test verisinde gerçekleşen tahsilat miktarı: 100.385.214

Klasik model tahmini tahsilat miktarı : 126.923.604

Yeni model tahmini tahsilat miktarı : 109.686.969

5. Sonuç

Finans sektöründe önemli bir kriter olan temerrüt halinde kayıp tahmini. Borçlu müşterilerin borçlarını ödememeleri durumunda bankaların karşılaşılabileceği risk tutarını belirlemek için kullanılmaktadır. Genel olarak temerrüt tarihinden itibaren dosyanın kapanma yani tahsilat imkânı kalmayana kadar gelişen süreçte oluşan bir kümülatif tahsilat oranı eğrisi tahmin edilmektedir. Bu eğriler her bir müşteri için kümülatif zaman serisi iken bir portföy için kesit veridir. Bu sebeple panel veri özelliği taşıyan temerrüt halinde kayıp modellerinin kurulması zorlaşmaktadır.

Klasik yaklaşımla kümülatif eğrinin son noktası olan toplam kayıp miktarını karar ağaçları yöntemi ile homojen sınıflara ayırırken 2. aşamada her bir sınıfın içinde ye alan müşterilerin kümülatif zaman noktalarındaki tahsilat oranlarının ortalaması ilgili sınıfın beklenen tahsilat eğrisi olarak kabul edilir.

Bu tezde önerilen yaklaşımda ise homojen olması istenen ilgili sınıflamanın eğrilerin son noktasına göre değil eğrilerin tamamı için oluşturulması fikrine dayanmaktadır. Bu sebeple karar ağaçları yöntemine geçmeden önce panel veri durumundaki bağımlı değişken kümeleme algoritmaları ile homojen gruplara ayrıldıktan sonra her bir grup bir bağımlı değişken kategorisi haline getirilerek karar ağacı yöntemi ile tahmin edilir.

Kümeleme yöntemi ile oluşturulan gruplarda tüm tahsilat eğrisi dikkate alındığından tahmin edilen grupların ortalama tahsilat eğrişine göre grup gözlemlerinin tahsilat eğrisi varyansları da klasik yöntemde göre düşük çıkması beklenmelidir. Ancak önerilen yöntemin özellikle toplan tahsilat tutarını tahmin açısından sürekli seri olan bağımlı değişkeni kategorilere indirilmesi veri kaybında beraberinde getirecektir.

Ortalama kare hata tüm zaman noktalarında yeni yöntemde klasik yöntemde göre daha düşük çıkmıştır. Tüm tahsilat zaman noktalarındaki hataların ortalamasında klasik yöntemde 36,72 iken yeni yöntemde 29,23 olmuştur.

Beklenen toplam tahsilat tutarlarının test verisinde doğru tahmin edilip edilmediđi incelendiđinde test verisinde 100 milyon olarak gerekleřen tahsilat tutarı. Klasik modelde yaklaşık 126 milyon olarak tahmin edilirken yeni yöntemde yaklaşık 109 milyon olarak daha yakın tahmin elde edilmiřtir.

Bu sonuçlar dođrultusunda temerrüt halinde kayıp tutarı ve tahsilat eđrilerinin tahmin edilmesinde tek başına karar ađacı yerine kümeleme algoritmaları ile karar ađaçlarının birlikte kullanılması tahmin hatalarını ve varyansını azaltmaktadır.



6. Kaynaklar

- Araten, M; Jacobs, M Jr; Varshney, P** (2004), ‘Measuring G on Commercial oans: An 18-Year Internal Study’, The RMA Journal, May 2004
- AY, A., ERDOĞAN, S. ve UYSAL, Doğan** (2007), “İstikrar Politikaları ve Türkiye IMF İlişkileri”, Türkiye Ekonomisi, (ed) Çizgi Kitabevi, Konya, ss. 579- 613.
- Aydın, T.**, (2020)TFRS 9 Standardı Kapsamında Karşılık Uygulamalarının Türk Bankacılık Sektörüne Etkisinin İncelenmesi
- BDDK**, TFRS 9 Uyarınca Beklenen Kredi Zararı Karşılığı Hesaplamasına İlişkin Rehber, https://www.bddk.org.tr/ContentBddk/dokuman/mevzuat_0085.pdf
- BIS**,”Basel II International Convergence of Capital Measurement and Capital Standards: a Revised Framework” 10 Jun 2004
- BIS**,”Basel III: BIS,“Revisions to the standardised approach for credit risk” 22 12 2014
- Dermine, J and Carvalho, C** (2005), ‘Bank loan losses-given-default: A case study’, Journal of Banking & Finance, 30 pp. 1219- 1243
- EUR-Lex**,Consolidated text: Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms and amending Regulation (EU) No 648/2012 (Text with EEA relevance)Text with EEA relevance
- E.B.Authority**, EBA/GL/2017/16 ”Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures” 23/04/2018
- Frye, J** (2000), ‘Depressing Recoveries’, Risk, 13(11) pp. 108–111.
- Gupton, G** (2005), ‘Advancing Loss Given Default Prediction Models: How the Quiet Have Quickened’, Economic Notes by Banca Monte dei Paschi di Siena SpA, (34-2) pp. 185–230
- Günel, M.**, (2006), Para Banka Ve Finansal Sistem
- Gürtler, M., Hibbeln, M.**, 2013. Improvements in loss given default forecasts for bank loans. Journal of Banking & Finance 37, 2354–2366.
- Kalkan, A.**, 2007, Bankacılıkta Risk Yönetimi ve Türk Bankacılık Sistemi Üzerine Bir Değerlendirme, ss. 20
- Kaya, F.**, 2012, Bankacılık Giriş ve İlkeleri.

- Khieu, H; Mullineaux, ; Yi, H** (2012), ‘The determinants of bank loan recovery rates’, Journal of Banking & Finance, 36 pp. 923-933
- Leow, M., Mues, C. & Thomas, L. C.** (2009) Modelling Loss Given Default for Mortgage Loans, Presentation in Conference of Credit Scoring and Credit Control XI, Edinburgh, August 2009.
- Resmi Gazete** no:29511 “Bankaların sermaye yeterliliğinin ölçülmesine ve değerlendirilmesine ilişkin yönetmelik” 23/10/2015
- Resmi Gazete** no:29511 “Kredi riskine esas tutarın içsel derecelendirmeye dayalı yaklaşımlar ile hesaplanmasına ilişkin tebliğ” 23/10/2015
- Resmî Gazete,** 23/10/2015, Sayı: 29511, Kredi Riskine Esas Tutarın İçsel Derecelendirmeye Dayalı Yaklaşımlar İle Hesaplanmasına İlişkin Tebliğ
- Saltoğlu, B.,** 2009, Finansal Risk Yönetimi, İstanbul, ss. 37
- Schuermann, T** (2004), ‘What Do We Know About Loss Given Default?’, Working Paper, (accessed 2014-03-18) Available from: <http://ssrn.com/abstract=525702>
- Qi, M. & Yang, X.** (2009) Loss given default of high loan-to-value residential mortgages, Journal of Banking & Finance, 33, 788-799
- TCMB,** Türkiye’nin Güçlü Ekonomiye Geçiş Programı 2001, <https://www.tcmb.gov.tr/wps/wcm/connect/26640b7b-9641-4c35-99ec-cd10a9d4e51b/program.pdf?MOD=AJPERES&CACHEID=ROOTWOR KSPACE-26640b7b-9641-4c35-99ec-cd10a9d4e51b-m3fB7oF>