

**T.C.  
DOKUZ EYLÜL ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
İŞLETME ANABİLİM DALI  
İŞLETME PROGRAMI  
DOKTORA TEZİ**

**TÜKETİCİ KREDİ RİSKLERİNİN, META-SEZGİSEL  
YAKLAŞIMLAR İLE  
İYİLEŞTİRİLMİŞ RASTGELE ORMANLAR YÖNTEMİ  
ARACILIĞIYLA DEĞERLENDİRİLMESİ**

**Hazar ALTINBAŞ**

**Danışman  
Prof. Dr. Gökтуğ Cenk AKKAYA**

**İZMİR- 2019**

**DOKTORA**  
**TEZ ONAY SAYFASI**

**Üniversite** : Dokuz Eylül Üniversitesi  
**Enstitü** : Sosyal Bilimler Enstitüsü  
**Adı ve Soyadı** : HAZAR ALTINBAŞ  
**Öğrenci No** : 2012800682  
**Tez Başlığı** : Tüketici Kredi Risklerinin, Meta-Sezgisel Yaklaşımlar ile İyileştirilmiş Rastgele Ormanlar Yöntemi Aracılığıyla Değerlendirilmesi

**Savunma Tarihi** : 14/01/2019

**Danışmanı** : Prof.Dr.Göktuğ Cenk AKKAYA

**JÜRİ ÜYELERİ**

<u>Ünvanı, Adı, Soyadı</u>	<u>Üniversitesi</u>	<u>İmza</u>
Prof.Dr.Göktuğ Cenk AKKAYA	DOKUZ EYLÜL ÜNİVERSİTESİ	
Prof.Dr.Ali ÖZDEMİR	DOKUZ EYLÜL ÜNİVERSİTESİ	
Prof.Dr.Ahmet ÖZEN	DOKUZ EYLÜL ÜNİVERSİTESİ	
Prof.Dr.Türker SUSMUŞ	EGE ÜNİVERSİTESİ	
Doç.Dr.Emre GÜLER	İZMİR KATİP ÇELEBİ ÜNİVERSİTESİ	

HAZAR ALTINBAŞ tarafından hazırlanmış ve sunulmuş olan bu tez savunmada başarılı bulunarak oy birliği / oy çokluğu ( ) ile kabul edilmiştir.

**Prof. Dr. Metin ARIKAN**  
Müdür

## YEMİN METNİ

Doktora Tezi olarak sunduđum ‘‘Tüketici Kredi Risklerinin, Meta-Sezgisel Yaklaşımlar ile İyileştirilmiş Rastgele Ormanlar Yöntemi Aracılığıyla Deđerlendirilmesi’’ adlı çalışmanın, tarafımdan, akademik kurallara ve etik deđerlere uygun olarak yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuđunu, bunlara atıf yapılarak yararlanılmış olduđunu belirtir ve bunu onurumla dođrularım.



Tarih

22/01/2019

Adı SOYADI

Hazar ALTINBAŞ

İmza

## ÖZET

Doktora Tezi

Tüketici Kredi Risklerinin, Meta-sezgisel Yaklaşımlar ile  
İyileştirilmiş Rastgele Ormanlar Yöntemi Aracılığıyla Değerlendirilmesi

Hazar ALTINBAŞ

Dokuz Eylül Üniversitesi

Sosyal Bilimler Enstitüsü

İşletme Anabilim Dalı

İşletme Programı

2007-2009 yıllarına dayanan küresel finans krizi, kredi risklerinin değerlendirilmesi analizlerine yönelik artan bir ilgiyi de beraberinde getirmiştir. Geri ödeme gücü zayıf çok sayıda hane halkına verilen konut kredilerini, bir diğer deyişle eşik altı kredileri ipoteğe dayalı türev ürünlere çevirerek başka yatırımcılara satan köklü ve büyük bankalar, birer sistemik risk kaynağı haline gelmiş ve maruz kaldıkları temerrüt risklerinin bütün finansal sisteme yayılmasına sebep olmuşlardır.

Kredi talebinde bulunanlara ilişkin bilgi toplama ve değerlendirme yapmaya yönelik modern çalışmaların geçmişi 1940'lı yılları bulsa da istatistiksel yöntemlerin kullanımına ilişkin ilk örnekler 1960'ların sonuna doğru ortaya çıkmıştır. O zamanlardan günümüze kadar, hızla artan kredi miktarları ve kredi kullanan kişi sayısı ile paralel olarak borçlananlar ile ilgili daha fazla veri depolanmaya başlanmıştır. Bu verileri analiz edecek ve yoğun matematiksel hesaplamaya dayalı gelişmiş yöntemler de bilgisayar bilimi ile birlikte gelişme göstermiştir. İhtiyaçların ve çözümlerin bir sarmal şeklinde etkileşerek bugünlere gelmesiyle birlikte de kredi başvurularını değerlendirme yöntemlerine ilişkin geniş bir yazın ortaya çıkmıştır.

Yazının çok büyük bir bölümünü istatistiksel ve makine öğrenme yöntemleri oluşturmaktadır. Bu yöntemlerin birçoğu, çok fazla değişken ve gözlem içeren karmaşık veri kümelerinden yorumlanabilir bilgi ve doğru tahmin ortaya çıkartmakta oldukça başarılı olmaktadır. Ancak yöntem başarılarında

ortaya çıkan çok küçük farklılıklar şirket kar veya zararlarında ve hatta sistemin sürdürülebilirliği üzerinde çok büyük etkiler yaratabilecektir. Bu nedenle mevcut yöntem ve yaklaşımları iyileştirme ve geliştirme çabaları halen sürmektedir.

Söz konusu çabaların bir parçası olarak bu çalışmada, performans kıyaslamasına imkân vermesi sebebiyle birçok çalışmada incelenmiş olan bir tüketici kredi veri kümesi hibrit bir yaklaşım ile analiz edilmiştir. Veri kümesinin öğrenimi ve tahminler, biyoloji ve tıp gibi alanlarda yapılan analiz çalışmalarında yoğun olarak kullanılmakla birlikte finans alanında kendine yeni yeni yer bulmaya başlayan bir makine öğrenme yöntemi olan Rastgele Ormanlar ile gerçekleştirilmiştir. Yöntemin iyileştirilmesi amacıyla da Genetik Algoritma ve Tavlama Benzetimi meta-sezgisel yaklaşımları ile değişken seçimi yapılmıştır. Böylelikle yöntemin eğitim sürecini olumsuz etkileyerek tahmin performansını düşüren değişkenlerin analizlerden elenmesi amaçlanmıştır.

Tez ile önerilen bir dizayn ile birleştirilen meta-sezgisel yaklaşımların içerisinde Rastgele Ormanlar yöntemi gömülü olarak kullanılmıştır. Birleştirilmiş dizayn ile meta-sezgisel yaklaşımların güçlü yönleri kullanılarak zayıflıklarının bertaraf edilmesi sağlanmıştır. Kredi değerlendirme çalışmalarında, Genetik Algoritma sıklıkla kullanılmakla birlikte Tavlama Benzetimi uygulama örneği bulunmamaktadır.

Sonuçlar Rastgele Ormanlar yönteminin hibritleştirilerek kullanımında, tek başına kullanımına göre tahmin performanslarında istatistiksel olarak anlamlı artışlar sağlandığını göstermektedir. Aynı veri kümesi ile analiz yapılan yazındaki diğer çalışmalar ile karşılaştırıldığında da önerilen hibrit yöntemin yüksek bir performans gösterdiği görülmektedir. Yöntem kredi başvuru değerlendirmelerinde karar vericilere büyük faydalar sağlayacaktır.

**Anahtar kelimeler:** Kredi Risk Değerlendirmesi, İstatistiksel Öğrenme, Makine Öğrenme, Rastgele Ormanlar, Meta-sezgiseller, Değişken Seçimi.

## **ABSTRACT**

**Doctoral Thesis**

**Doctor of Philosophy (PhD)**

**Assessment of Consumer Credit Risk via**

**Random Forests Method Improved with a Combined Meta-heuristic Approach**

**Hazar ALTINBAŞ**

**Dokuz Eylül University**

**Graduate of Social Sciences**

**Department of Business Administration**

**Business Administration Program**

**Global financial crisis, which dates back to 2007-2009, raised a growing interest to credit risk assessment analyses. Long-established and big banks became a source of systemic risk and caused their exposure to debt risks to be transmitted throughout the whole financial system by landing mortgages to too many households with low repayment capabilities, which are also known as subprime mortgages.**

**History of modern studies for collecting information and assessment about credit applicants have begun around 1940s but first examples of statistical methods usage showed up in late 1960s. From that day on, in parallel to rapidly increasing amount of credit and credit borrowers, much more data has been collected. Along with progress made in computer science, advanced methods which are based on intensive computation to analyze these data are also developed. Needs and solutions are mutually interacted in a spiral way up to the present and a wide literature has been emerged about credit applications' assessment.**

**Most of the literature is consisted of statistical and machine learning methods. Most of these methods are very successful for deriving interpretable information and correct predictions from complex datasets with many variables and observations. However, little differences will have great impacts on company**

profits or losses and even on the sustainability of the system. That's the reason for continuing efforts to improve and develop existing methods and approaches.

As a part of these efforts, in this study, a well-known consumer credit dataset which is used as a benchmark set for comparison purposes, is analyzed with a hybrid approach. Dataset learning and predictions are conducted with a machine learning method called Random Forests, which has been used intensely in subjects like biology and medicine but became used in finance recently. In order to improve the method, feature selection is applied with Genetic Algorithm and Simulated Annealing meta-heuristic approaches. Aim is to eliminate irrelevant variables that disturb learning process and cause worse predictions.

Random Forests method is embedded to the design of combined meta-heuristics proposed by this thesis. Combined design allowed to use strengths of meta-heuristics while overcoming weaknesses. In credit assessment studies, Genetic Algorithm is widely used but there is no example of a study with Simulated Annealing.

Results show that hybridized implementation of Random Forests produces statistically significant increases in prediction performances vis a vis sole usage. When compared with other studies adopted the same dataset, it is seen that proposed hybrid method shows high performance. Method will be highly beneficial for decision makers in assessment of credit applications.

**Keywords:** Credit Risk Assessment, Statistical Learning, Machine Learning, Random Forests, Meta-heuristics, Feature Selection.

**TÜKETİCİ KREDİ RİSKLERİNİN, META-SEZGİSEL  
YAKLAŞIMLAR İLE İYİLEŞTİRİLMİŞ RASTGELE ORMANLAR  
YÖNTEMİ ARACILIĞIYLA DEĞERLENDİRİLMESİ**

**İÇİNDEKİLER**

TEZ ONAY SAYFASI	ii
YEMİN METNİ	iii
ÖZET	iv
ABSTRACT	vi
İÇİNDEKİLER	viii
KISALTMALAR	xi
TABLOLAR LİSTESİ	xiii
ŞEKİLLER LİSTESİ	xiv
EKLER LİSTESİ	xv
GİRİŞ	1

**BİRİNCİ BÖLÜM**

**KREDİ DEĞERLENDİRMESİNE YÖNELİK ULUSLARARASI  
DÜZENLEMELER VE ANALİZ YAKLAŞIMLARI**

1.1. FİNANSAL MİMARİ	4
1.1.1. Finansal Krizler	7
1.1.2. Basel I ve Basel II	11
1.1.3. Yeni Finansal Mimari ve Eleştiriler	16
1.1.4. Basel III	18
1.2. KREDİ RİSK DEĞERLENDİRMESİ	21
1.3. İSTATİSTİKSEL VE MAKİNE ÖĞRENME YÖNTEMLERİ	25
1.3.1. Parametrik Yöntemler	27
1.3.2. Parametrik Olmayan Yöntemler	29



1.3.3.	Rastgele Ormanlar Yönteminin Tercih Edildiği Çalışma Konuları	33
1.4.	META-SEZGİSEL ALGORİTMALAR	34
1.4.1.	Biyo/doğa-ilhamlı Hesaplama	35
1.4.2.	Meta-sezgisel Algoritmaların Kredi Değerlendirme Uygulamaları	36
1.4.3.	Meta-sezgisel Algoritmaların Diğer Finans ve Mühendislik Alan Uygulamaları	38

## İKİNCİ BÖLÜM

### TÜKETİCİ KREDİ BAŞVURULARININ DEĞERLENDİRİLMESİNDE RASSAL ORMANLAR YÖNTEMİ VE META-SEZGİSEL ALGORİTMALAR İLE İYİLEŞTİRME ÖNERİSİ

2.1.	AĞAÇ TABANLI YÖNTEMLER	40
2.1.1.	Karar Ağaçları	41
2.1.1.1.	Ağaç Budaması	45
2.1.1.2.	Torbalanmış Ağaçlar	46
2.1.2.	Rastgele Ormanlar	46
2.1.2.1.	Gini Endeksi Örnek Bir Ayrım Kararı	51
2.2.	GENETİK ALGORİTMA- TAVLAMA BENZETİMİ	54
2.2.1.	Değişken Seçimi	55
2.2.2.	P, NP ve NP-zor Problemler	57
2.2.3.	Genetik Algoritma	59
2.2.3.1.	Kromozom/birey	59
2.2.3.2.	Popülasyon	60
2.2.3.3.	Seçim	61
2.2.3.4.	Yeniden Oluşturma	62
2.2.3.5.	Mutasyon	63
2.2.3.6.	Yeni Popülasyon Oluşturma ve Genetik Algoritma Adımları	64
2.2.4.	Tavlama Benzetimi	65
2.2.4.1.	Çözüm Gösterimi ve Komşuluk Oluşturma	67
2.2.4.2.	Simülasyon Bileşenleri ve Adımları	68

2.2.5.	Birleştirilmiş Algoritmalar- Rastgele Ormanlar Dizaynı	70
2.2.6.	Genetik Algoritma ve Tavlama Benzetimi Parametre Değer ve Seçimleri	73
2.3.	PERFORMANS DEĞERLENDİRME ÖLÇÜTLERİ	74

## ÜÇÜNCÜ BÖLÜM

### ÖRNEK KREDİ DEĞERLENDİRME VERİ KÜMESİ ÜZERİNDE BİR İNCELEME

3.1.	İNCELEMENİN AMACI	78
3.2.	ÖNERİLEN ANALİZ YÖNTEMİNİN GENEL AKIŞI	80
3.3.	RASTGELE ORMANLAR BULGULARI	84
3.4.	BA-RO BULGULARI	86
3.5.	ANALİZ BULGULARININ KARŞILAŞTIRILMASI	87
3.6.	TARTIŞMA	88
	SONUÇ	96
	KAYNAKÇA	105
	EKLER	



<b>NP</b>	Non-deterministic Polynomial- Belirleyici Olmayan Polinomsal
<b>OECD</b>	The Organization for Economic Co-operation and Development- Ekonomik İş birliđi ve Kalkınma Örgütü
<b>P</b>	Polynomial- Polinomsal
<b>PSO</b>	Parçacık Sürü Optimizasyonu
<b>RO</b>	Rastgele Ormanlar
<b>s.</b>	Sayfa No
<b>ss.</b>	Sayfadan Sayfaya
<b>SA</b>	Sinir Ağları (Yapay Sinir Ağları kullanımını da kapsamaktadır)
<b>SYO</b>	Sermaye Yeterlilik Oranı
<b>TB</b>	Tavlama Benzetimi
<b>TBA</b>	Temel Bileşenler Analizi
<b>WB</b>	World Bank- Dünya Bankası
<b>YN</b>	Yanlış Negatif
<b>YP</b>	Yanlış Pozitif

## TABLULAR LİSTESİ

Tablo 1: Rastgele ormanlar yönteminin kullanım konu ve amaçları .....	33
Tablo 2: Meta-sezgisel algoritma kullanım örnekleri .....	39
Tablo 3: Dört yeni mağaza için çocuk koltuğu satışı tahmin değişkenleri .....	50
Tablo 4: Dört yeni mağaza için çocuk koltuğu satışı tahminleri .....	50
Tablo 5: Genetik algoritma parametreleri .....	73
Tablo 6: Tavlama benzetimi parametreleri .....	74
Tablo 7: Kredi sınıflandırması için karışıklık matrisi .....	74
Tablo 8: Çalışmada kullanılan veri kümesine ait değişken bilgileri ve kategorik özellikleri.....	79
Tablo 9: RO yöntemi ile kredi riski tahmin sınıflandırma bulguları .....	85
Tablo 10: BA-RO yöntemi ile kredi riski tahmin sınıflandırma bulguları.....	87
Tablo 11: Wilcoxon testi sonuçları .....	88
Tablo 12: Farklı test kümesi büyüklüklerinde seçilen değişken sayısı ve genel ortalaması .....	90
Tablo 13: Farklı test kümesi büyüklüklerinde değişkenlerin seçim yüzdesi ve genel ortalamalar.....	91

## ŞEKİLLER LİSTESİ

Şekil 1: Karar ağacı bileşenleri .....	41
Şekil 2: Bir karar ağacı örneği .....	42
Şekil 3: İki boyutlu veri uzayında KA ayırması ve karar bölgelerinin oluşumu .....	43
Şekil 4: Yanlılık ve varyansın grafiksel gösterimi .....	44
Şekil 5: Rastgele Ormanlar ile çocuk koltuğu satışı analizi örneği .....	48
Şekil 6: Karar Ağaçları ile çocuk koltuğu satışı analizi örneği.....	49
Şekil 7: Örnek ayırma kararı süreci .....	52
Şekil 8: İkili olarak kodlanmış bir birey örneği .....	59
Şekil 9: Tek nokta çaprazlama ile yeni bireylerin oluşturulması.....	63
Şekil 10: İkili kodlanmış bireylerde mutasyon .....	64
Şekil 11: Genetik algoritma tek çevrim akışı.....	65
Şekil 12: Tavlama benzetimi çözüm oluşturma süreci .....	68
Şekil 13: Veri kümesinin analiz aşamaları: birinci aşama .....	81
Şekil 14: Veri kümesinin analiz aşamaları: ikinci aşama.....	82
Şekil 15: Veri kümesinin analiz aşamaları: üçüncü aşama .....	83
Şekil 16: Müşterilerin vadesiz mevduat hesap durumları ve sınıf dağılımları.....	92
Şekil 17: Müşteri kredilerinin vadesi ve sınıf dağılımları.....	93
Şekil 18: Müşterilerin kredi geçmişleri ve sınıf dağılımları .....	93
Şekil 19: Müşterilerin çektikleri kredi miktarı ve sınıf dağılımları .....	94
Şekil 20: Müşterilerin tasarruf hesap ve tahvil durumları ve sınıf dağılımları .....	94
Şekil 21: Müşterilerin iş durumları ve sınıf dağılımları.....	95

## **EKLER LİSTESİ**

- EK 1: Yirminci Yüzyılda Yaşanmış Ekonomik ve Finansal Krizler Listesi ..... ek s.1  
EK 2: Alman Tüketici Kredi Veri Kümesi, Gözlemlere Ait Değişkenler ..... ek s.2  
EK 3: Değişkenlerin Betimleyici İstatistikleri ..... ek s.6



## GİRİŞ

Kredi, güven anlamına gelmektedir. Sosyal alanda ve sosyal bir parçası olarak özellikle ekonomik alanda kredi, insan birlikteliğinin ve iş birliğinin olduğu her yerde var olur. Nitekim insan grupları içi ya da arası güvene dayalı ilişkiler, tarihsel süreç boyunca büyüyen ve bugün bir parçası olduğumuz ve milyarlarca insanı kapsayan muazzam bir ağı ortaya çıkarmıştır.

Modern anlamda finansal sistemin izleri, Batı Avrupa'da feodal ekonomik yapılanmanın sonunu getiren merkantilist döneme kadar sürülebilir. M.S. 10-11. yy. sonrası yıllar, en güçlü ve etkili uygulayıcısı Venedik şehir devleti olan; kurumlar ve yasalar aracılığı ile özel mülkiyetin, ticaretin, sermaye birikim ve aktarımlarında önemli gelişmelerin yaşandığı ve gerek kıta içi gerekse kıtalar arası ilişkilerin ve etkileşimlerin yükselişe geçtiği bir dönemdir. Bu yükseliş, M.S. 15. yüzyıldan sonra, yine Batı Avrupalıların öncüsü olduğu coğrafi keşiflerin kaynak zeminini oluşturmuş, gelişmiş kredi sistemleri ile desteklenmeye devam ederek tarihte görülmemiş bir ekonomik büyüme yaratmıştır. Tüm dünyada, gayrisafi yurt içi hasıla (GSYH) yıllık ortalama bileşik büyüme oranları, M.S. 0- 1000 arası %0.01, 1000- 1820 arası %0.22, 1820- 1998 arası %2.21 olmuştur (Maddison, 2001: 28). Bu değişimde, gelişen bankacılık sistemi ve büyüyen ve çeşitlenen fon araçlarının etkisi olduğunu söylemek mümkündür.

Bugün gelinen noktada kredi hareketleri, ulusal sınırların ötesinde var olmakta, ulusları, şirketleri ve bireyleri benzer pazarlar üzerinden birbirine bağlamakta ve bir anlamda da bağımlı hale getirmektedir. 2015 yılında, Uluslararası Ödemeler Bankası'na (Bank of International Settlements- BIS) bilgi sağlayan tüm ekonomilerde, finans dışı sektörlere verilen toplam kredi miktarı, GSYH'lerinin %231.7'si kadar olmuştur (BIS, 2017a). Artan iletişim imkanları, ilişkiler ve geliştirilen mekanizmalar, krediler ve kredilerin kaynak yaratma potansiyelleri üzerindeki sınırları ortadan kaldırmıştır. Ancak bu kredilerin, yapıları gereği kaçınılmaz olarak çeşitli riskler taşımakta olduğunu da göz ardı etmemek gerekir.

Kredi hareketleri ile ortaya çıkan riskler, kredi şoklarına sebep olabilmektedir. Kredi şokları, piyasada mevcut kredi miktarlarında ani düşüşler ve/veya borçlanma maliyetlerinde hızlı yükselişler sonucu ortaya çıkabilir. Bunlara, piyasalarda



spekülasyonun yoğunlaşması, eşik altı kredilerde artışlar, düzenleyici ve denetleyici faaliyetlerin eksikliği, hatalı politikalar ve düşük düzeylerde seyreden faiz oranları gibi sebepler neden olmaktadır. Kredi şoklarına en güncel örnek olarak, küresel ölçekte sonuçları olan ve 2007 yılı sonlarına dayanan eşik altı ve ipoteğe dayalı tüketici kredisi krizi, bir diğer adıyla küresel finansal kriz gösterilebilir. Eşik altı olarak adlandırılan ve aldığı krediyi geri ödeme kapasitesi yeterli olmayan çok sayıda tüketiciye verilen konut kredileri, geri ödememelerin başlaması ile birlikte önce Amerika Birleşik Devletleri'nde (ABD), uluslararası finansal kuruluşlar ve piyasada yaşanan panik aracılığıyla da çok kısa süre içerisinde tüm dünyada büyük ekonomik problemlere sebep olmuştur. Kriz öncesi dönemde, yıllık Dünya GSYH büyüme oranları (WB, 2017a) ve işsizlik oranları (WB, 2017b), sırasıyla, 2006 yılında %4.326 ve %5.947, 2007 yılında %4.256 ve %5.54 iken, negatif etkilerin yoğunlaştığı 2008 yılında %1.819 ve %5.683, 2009 yılında ise -%1.735 ve %6.217 olmuştur.

Küresel finansal kriz sonrası, tüketici kredilerinin doğru değerlendirilmemesi ve kontrolsüz artışı sonucu ortaya çıkan durum, kredi değerlendirme konusuna hızla artan bir ilgiyi de beraberinde getirmiştir. Kredi şoklarının, sistemik bir risk unsuru olarak hızlı ve engellenemez yayılışı, risklerin en aza indirilmesi açısından önleyici bir yaklaşım olarak kredi başvurusunda bulunanın kredibilitesinin doğru belirlenmesinin önemini ortaya koymaktadır. Finansal mimariyi sağlamlaştırarak finansal istikrarı artırmak amacıyla, Uluslararası Para Fonu (International Monetary Fund- IMF), Dünya Bankası (World Bank- WB), Avrupa Merkez Bankası (European Central Bank- ECB) ve BASEL komitesi gibi uluslararası ölçekte düzenleyici kuruluşlar çeşitli önlemler almıştır. Bu önlemler içerisinde, kredi risklerinin doğru değerlendirilerek sermaye yeterliliklerinin gereken düzeylere çekilmesi de yer almaktadır. Sermaye tabanının kalitesinin artırılmasının yanı sıra, tüm risk unsurlarının sermaye yeterliliği çerçevesi tarafından başarılı şekilde ele alınması gerekmedir (BIS, 2010a).

Tüketici kredi risklerinin değerlendirilmesi, rekabetçi ortamın da etkisiyle yetersiz analiz ve sübjektif yargılara bağlı kalabilmektedir. Bu eksiklikleri gidermek adına, teknolojik ilerleme sonucu muazzam bir artış gösteren işlem gücünden yararlanarak, büyük miktarlarda veriyi hızlı ve doğru şekilde analiz etme ve bilgiye dönüştürme şansı bulunmaktadır. Makine ve istatistiksel öğrenme, meta-sezgisel algoritmalar gibi veri madenciliği yöntemleri, karar verme noktalarında başarılı

şekilde kullanılabilir. Kredi başvurusunda bulunanları, temerrüde düşmeyen (“iyi”) ve temerrüde düşen (“kötü”) olarak sınıflandıran bu yöntemler, kredi verenlerin karar süreçlerine destek olarak hem rekabet edebilirliklerini artırmakta hem de daha düşük düzeyde riske maruz kalmalarını sağlamaktadır. Azalan risk de sonuç olarak daha sağlıklı bir finansal sistemin oluşmasını sağlayacaktır.

Hem akademik hem de pratik çalışmalarda kullanılan sınıflandırma yöntemleri arasında, tek başına “en iyi” olarak adlandırılabilir biri bulunmamaktadır. Analiz edilen verinin niteliği gereği, farklı yöntemleri karşılaştırmak gerekebilir; ya da bir arada kullanarak daha başarılı sonuçlar elde etmek mümkün olabilir. Bu çalışmada, başarılı bir tahmin modeli olarak yazında sıkça yer alan ve gözetimli bir öğrenen yöntem olan rastgele ormanlar (RO) yöntemi ile bir tüketici kredi başvuru veri seti analiz edilmiştir. Söz konusu veri seti, kredi risk değerlendirme yöntemlerinin performanslarının değerlendirilmesi ve karşılaştırılması amacıyla analizlerde sıkça kullanılmaktadır. RO yönteminin performansını artırmak amacıyla iki farklı meta-sezgisel algoritmaya başvurulmuştur. Kompleks optimizasyon problemlerinin çözümünde sıklıkla başvuru yapılan Genetik Algoritma (GA) ile Tavlama Benzetimi (TB) algoritmaları bir arada kullanılarak, her iki yöntemin sahip olduğu güçlü taraflardan yararlanmak hedeflenmiştir. Bundan sonra birleştirilmiş algoritma (BA) olarak anılacak olan bu yöntem, veri seti üzerinde değişken/nitelik seçimi gerçekleştirerek daha anlamlı bir analiz yapılmasını sağlamıştır.

Tez, giriş ve sonuç hariç üç ana bölümden oluşmaktadır. Birinci bölümde, çalışmaya ilişkin teorik ve pratik çerçeve sunulmuştur. İkinci bölümde, analizde kullanılan yöntemler tanıtılmış, yeni yöntem önerisi açıklanmış ve parametre seçimleri verilmiştir. Üçüncü bölümde analiz bulguları sunulmuş ve mevcut yazın ile karşılaştırmalı olarak tartışılmıştır. Tez, genel bir değerlendirmenin yapıldığı ve ileride gerçekleştirilebilecek çalışmalara yönelik önerilerin sunulduğu sonuç bölümüyle son bulmaktadır.

# **BİRİNCİ BÖLÜM**

## **KREDİ DEĞERLENDİRMESİNE YÖNELİK ULUSLARARASI**

### **DÜZENLEMELER VE ANALİZ YAKLAŞIMLARI**

Gelişmiş ekonomiler başta olmak üzere dünyanın pek çok ekonomisini derinden etkileyen küresel finans krizi, hükümetlerin, ekonomik tarafların ve karar vericilerin mevcut iktisadi ve finansal teori ve stratejilerini gözden geçirmelerine neden olmuş, sorunlara cevap veremeyen ve hatta sorunlara sebep olduğu düşünülen yaklaşımlar ve çalışmalar tespit edilip terk edilmeye ya da iyileştirilmeye başlanmıştır. Yoğun şekilde entegre olmuş ekonomilerin büyük krizlerin önlenesi için birlikte çalışma ihtiyacı da kaçınılmaz olmuştur.

Küresel ölçekte daha sağlam ve sürdürülebilir bir finans sisteminin oluşturulması amacıyla yapılan tüm çalışmaları finansal mimari kavramı altında birleştirmek mümkündür. Küresel finans krizi ile herkes, finans piyasalarının, ülkelerin reel üretimlerinin ve ekonomik performanslarının altında sağlam bir zemin sağlayabildiği gibi tüm finansal ve reel sistemin durdurulamaz şekilde zarar görmesine de sebep olabileceğini görmüştür. Firmaların üretim yatırımlarının artması ile istihdamın ve büyümenin yükseltilmesi ancak ve ancak güvenilir, aşırı spekülasyon ortamlarının oluşmasına izin vermeyen ve sürdürülebilir bir finansal piyasa desteği ile sağlanabilir.

Bu hedeflere ulaşmak için atılan adımlara ilişkin teorik ve pratik çerçevenin sunulduğu bu bölümde finansal mimari kavramı ele alındıktan sonra, kredi risk değerlendirmesinin bu çerçeve içerisindeki yerine değinilmiş ve küresel ölçekte yapılan çalışmalardan bahsedilmiştir. Son olarak da risk değerlendirmesinde kullanılan ve tez çalışmasının analiz dayanağını oluşturan yöntemlere ilişkin yazın taraması sunulmuştur.

#### **1.1. FİNANSAL MİMARİ**

İnsanlık tarihinde modern anlamda ekonomik faaliyetin temellerini tarım devrimlerinde bulmak mümkündür. İnsanlar iş birliği ile arazileri ekip biçmeye

başlamış, buna bağlı olarak çeşitli büyüklüklerde topluluklar oluşturmuş ve yoğun olmamakla birlikte zaman zaman başka topluluklarla ticari ilişkiler de kurmuştur. Örgütlenme biçimleri ve sınıfsal ayrımlar tarihsel süreç içerisinde biçim değiştirmiş olsa da tarıma dayalı üretim biçimi, küresel ölçekte emperyalist eğilimlerin yaygınlaştığı, kaynak yaratma yöntemi olarak tarımsal üretim ve buna bağlı vergilendirme ile fetih yoluyla yağmanın yerini kredilerin almaya başladığı 1400'lü yıllara kadar geçerliğini korumuştur.

Hayvanlardan Tanrılara: Sapiens isimli popüler kitabında Harari (Harari, 2014), insanlık tarihinde önemli bir değişim aşaması olarak gördüğü kapitalist üretim biçiminin temellerini, krediler ile gerçekleştirilen ve batı toplumuna Sanayi Devrimi'ne giden yolu açan Avrupa tipi emperyalizme dayandırmaktadır. Girişimcilerin desteğiyle elde edilen fonlarla oluşturulan deniz filoları sayesinde yeni kıtalar keşfedilmiş, buralardan elde edilen zenginlikler yatırımcılara kazanç olarak aktarılmış ve sağlanan bu güvenle birlikte yeni keşifler için gereken kaynakların bulunmasını sağlayan kredi ortamları oluşturulmuştur.

Ancak bu güven ortamları, düzenleyici ve denetleyici faaliyetlerin eksikliğinden kaynaklanan problemler ve yüksek kazançlara duyulan risk iştahı ile birlikte daha o dönemlerde sarsıntılara uğramıştır. Hollanda'da 1636 yılında ortaya çıkan 'Lale Soğanı Balonu' ve Fransa'da 1700'lü yılların başında ortaya çıkan 'Mississippi Balonu', yakın zamanda tüm dünyayı etkisi altına alan kripto paralarda - başta aralarında en bilineni olan 'Bitcoin' olmak üzere- ortaya çıkan ve bilgi ve denetim eksikliğine bağlı olarak gelişen problemlere tarihsel örnek olarak gösterilmeye devam edebilmektedir.

Acemoglu ve Robinson (2012), batı uygarlığının gelişim sürecinde kurumsal yaklaşım farklılıklarının önemli bir dönüm noktası olduğunu belirtirler. Modern toplumun temeli olan ve İngiltere'de başlayan 'Sanayi Devrimi', geleneksel kurumların yerine piyasa ekonomisini oluşturacak ve düzenleyecek merkezileştirilmiş yeni kurumların ortaya çıkması ile gerçekleştirilebilmiştir. Bu süreç, tarihte eşî görülmemiş şekilde artan bir üretim ortamı oluşturmuş ve bu arz fazlasının tüketimine açık yeni pazarlar aranmıştır. Başlangıçta zorla, günümüzde ise görece demokratik olduğu söylenebilecek yöntemlerle ülkeler arası ticaret gelişmeye, pazarlar entegre hale gelmeye ve nihai olarak da küreselleşme adı verilen ve yalnızca mal ve

hizmetlerin değil, bunları üretecek kaynağı sağlayacak sermayenin de hareket alanını genişletecek ve kolaylaştıracak bir dizi gelişme yaşanmaya başlamıştır. 21. yy. ise, söz konusu sermaye hareketlerinin saniyeler içinde gerçekleşmesini sağlayan teknolojik yeniliklere sahne olmuştur. Bu kadar büyük ölçekte ve hızla değişen bir sistemin, beraberinde sorunlar getirmesi de kaçınılmazdır.

Günümüzde gelişmiş ve gelişmekte olan ülkelerin dahil olduğu ekonomik ve kültürel entegrasyon ortamını tanımlayan kavram küreselleşmedir. Küreselleşme sosyal, kültürel, siyasal ve ekonomik yönleri olan çok-boyutlu bir süreçtir. Küreselleşmenin önemli bir boyutu da değişimdir: ekonomiden siyasete, sosyal politikadan kültüre birçok alandaki değişimi ifade eder. Bir diğer boyutu da dünya ile bütünleşmedir: ticaretin ve para akışlarının daha çok uluslararası rekabeti esas alan dışa açık politikalar üzerinde yoğunlaşması ile gerçekleşir (Obadan, 2006)

Küreselleşme, insanlığın bilinen tarihine kıyasla bile fazlasıyla yeni bir kavramdır ve içinde bulunduğumuz dönemin belirgin özelliği olan hızlı iletişim ve işlem yapma kabiliyetinin getirdiği etkilerle de birlikte finansal sistem açısından çeşitli sorunlara sebep olabilmektedir. Genel kabul gören açıklamaya göre, finansal çalkantılar sistemik olarak finansal faaliyetlerin hızlı büyümesi ve hızlı hareketleri ile ortaya çıkmaktadır. 1970'lerin başında Bretton Woods sisteminin yıkılmasından sonra ve özellikle 1980'lerin başından itibaren küresel finansal işlemler ve piyasalar, başta üretim ve dış ticaret olmak üzere, diğer ekonomik faaliyetlere göre çok daha hızlı büyümüştür. Küreselleşme ile birlikte ortaya çıkan sorunlar beraberinde, ekonomist ve finansçıların halen etkili bir yanıt vermeyi başaramadıkları bazı sorular da getirmiştir: Küresel finansal bunalımlar ve çalkantılar neden vardır? Bu tür bunalımlara kimler nasıl müdahale edebilir? Müdahale merkezleri IMF gibi kurumlar mıdır, yoksa örneğin G20<sup>1</sup> ve Finansal İstikrar Kurulu (Financial Stability Board- FSB) gibi çok ülkeli gruplar mıdır? Yoksa çözüm finansal piyasalara mı bırakılmalıdır? Çözüm üretmede daha sıkı kurallar mı, yoksa piyasa serbestliği mi öne çıkmalıdır? Küresel bunalımların yaşandığı zamanlarda bu sorular daha da yoğunlaşmaktadır.

Bu kapsamda ele alınan finansal mimari kavramı da çoğunlukla, finansal kriz riskinin azaltılması ve krizlerin sistemik bir biçimde yayılmasının önüne geçilebilmesi

---

<sup>1</sup> G20: Gelişmekte olan ekonomiye sahip 20 ülkeyi ifade eder.

için yapılan çalışmaları içermektedir. Finansal mimari, mikro düzeyde bir değerlendirme yaparak gerek özel gerekse kamu işletmeleri için temel ilkelerin ve belirli durumlar karşısında en iyi uygulamaların belirlenebilmesi üzerinde odaklanmaktadır. Mikro düzeyde yapılan değerlendirmeler kapsamında, borç veren bankaların, borç alma talebinde bulunan kişi, şirket ya da kurumların geri ödeme kapasitelerini ortaya koymaları ve risk yönetim süreçlerini buna göre dizayn etmeleri de bulunmaktadır. Yeterli sayıda veriye ve yüksek işlem gücüne sahip teknolojik altyapıları bulunan bankaların yapması gereken, modern istatistik ve algoritma tekniklerinden doğru şekilde yararlanarak anlamlı bilgi üretebilmektedir. İstatistiksel/makine öğrenme ve meta-sezgisel algoritmalar bu çaba içerisinde son yıllarda öne çıkan uygulamalardır.

Finansal mimari kavramının ortaya çıkışında, tarih boyunca tecrübe edilen finansal/ekonomik krizleri önleme çabalarının şüphesiz büyük bir etkisi vardır. Bu sebeple finansal mimarinin ortaya çıkışı incelenirken konuya, finansal krizler ile başlamak gerekir.

### **1.1.1. Finansal Krizler**

Kriz kavramı iktisat yazınında çöküntü, bunalım, durgunluk, güç dönem ya da buhran gibi terimlere karşılık gelen, genel bir ifadeyle de ülke ekonomisini önemli derecede sarsacak ani ve beklenmedik bir durumda ortaya çıkan, genel olayların bir sonucu olarak oluşan, ekonominin mevcut konumunu ve geleceğini etkileyen ve genelde önlem alınmakta geç kalınan olumsuz bir durum olarak ele alınmaktadır (Yücel ve Kalyoncu, 2011).

Kriz sürecinde ekonomik yapı içinde toplam talebin düşmesi, yatırımların azalması, işsizliğin artması ve sonuçta refah seviyesinin düşmesi gibi çeşitli makroekonomik olumsuzluklar kendini göstermektedir. Bir başka ifadeyle ekonomik kriz, mal ve hizmet üretiminde ani düşüşler veya duruşlar ve bunun sonucu ortaya çıkan, kitlesel işsizlik, ücretlerin azalışları, insanların yaşam standartlarındaki gerilemeler olarak kendini gösteren bir durum olarak da tanımlanabilir. Ancak iktisat tarihine bakıldığında yaşanan krizlerin aslında beklenmedik durumlar olmadığını, sinyallerin ve göstergelerin önceden ortaya çıktığını söylemek mümkündür.

Dünya tarihi, ilk ticaretin yapıldığı dönemlerden itibaren ekonomilerde yaşanmış pek çok dönemsel iniş ve çıkışlara sahne olmuştur. Küçük büyük tüm krizler, dünyadaki ticaret ve piyasa ağının içerisinde yer alan her ülkeyi etkilemektedir. Kapitalist sistemin işleyişi içinde, nedenlerine bağlı olarak değişik kriz türleri ile karşılaşmaktadır. Bunlar, sermaye birikimindeki yoğunlaşma ve artışına bağlı olarak kâr oranlarının düşmesiyle gündeme gelen krizler, aşırı üretim artışı sonucu oluşan krizler ve talep yetersizliği sonucu oluşan krizler olarak kategorize edilebilirler. Bununla birlikte sistemin küreselleştiği son 30 yıllık dönemde krizin nedenleri farklılaşmakta ve daha çok finansal sistem krizlerine dönüşmektedir (EK 1’de, 20. yy. ’da yaşanmış ekonomik/finansal krizlerin kronolojik bir listesi verilmiştir.). İkinci dünya savaşından sonra kalkınma iktisatçıları arasında en çok tartışılan konulardan biri olan kalkınma finansmanında, eldeki kaynakların ve sermaye birikiminin ekonomik büyümeyi uarması için en etkili ve verimli bir şekilde nasıl değerlendirilebileceği açıklanmaya çalışılmaktadır. 1960’lı yıllar boyunca da tartışmalar daha çok yurtiçi finansal yapı ve araçların ekonomik büyümeye nasıl katkıda bulunabileceği üzerinde devam etmiştir. Bu çerçevede finansman ve ekonomik büyüme arasındaki ilişkiyi açıklamak için daha çok finansal aracılık sürecine vurgu yapılmaktadır. Patrick (1966) finansal sistem ile ekonomik büyüme ilişkisindeki nedenselliği, arzın öncülük ettiği ve talebin takip ettiği görüşleri ile açıklamaya çalışmıştır. Arzın öncülük ettiği durumda yeni finansal kurumların yaratıldığı, finansal varlıkların arzının ve enstrümanlarının geliştiği ve bu gelişim sürecinin ekonomik büyümeyi uyardığı bir nedensellikten söz edilmektedir. Talebin takip ettiği yaklaşımda ise reel sektör büyümek için finansal sektöre yönelmekte, bu yönelim reel sektördeki çıktının artmasını, buna bağlı olarak ticaretin yaygınlaşmasını ve istihdamın artmasını sağlamaktadır. Burada asıl vurgulanmak istenen, reel sektörün büyümesiyle ekonomilerde finansal sisteme ve sistemin sağladığı hizmetlere olan talebin artacağı ve artan bu talep neticesinde finansal sistemin gelişeceğidir. Diğer bir ifadeyle, milli gelirdeki artış ne kadar hızlı olursa, girişimcilerin kaynak ihtiyaçları da finansal kurumların hizmetlerine talepleri de o kadar büyük olacaktır.

1970’li yıllarda ise tartışmalar, finansal sektör politikaları üzerine yoğunlaşmaktadır. Özellikle artan enflasyonun sonucu olarak ortaya çıkan negatif reel faiz olgusunun tasarruflar üzerinde olumsuz etkiler yaratacağı ve dolayısıyla yatırımlar

için gerekli olan fonların oluşmayacağı, bunun da ekonomik büyümeyi engelleyeceği vurgulanmaktadır. Devlet müdahalelerinden tam olarak bağımsız olmayan sektörün, değişen koşullara göre kendini hızla düzenleyebilmesi için sektördeki sınırlamaların kaldırılması ve sektör üzerindeki devlet denetimi ve müdahalesinin azaltılarak daha fazla özgürlük alanına kavuşturulması gerekliliği savunulmuştur.

Bu sürecin devamı olarak 80’li ve 90’lı yıllar, gelişmekte olan ülkelerin gelişmiş ekonomilere daha fazla entegre olduğu ve dışa açılma eğilimlerinin ortaya çıktığı bir dönem olmuştur. Bu eğilim, finansal açıdan özellikle yabancı sermayenin ülkelere girişlerinde engellerin kaldırılması, bürokratik işlemlerin kolaylaştırılması ve maliyetlerin düşürülmesi ile hızlanmıştır. Bunun yanı sıra entegrasyon aşamasındaki ülkelerde şeffaflık, hesap verebilirlik, güven ve demokratikleşme eğilimleri de sürece katkı sağlamıştır. Ancak pek çok gelişmekte olan ülkede, makroekonomik istikrarsızlıklar giderilmeden ve kültürel farklılıklar göz önüne alınmadan içsel ve dışsal finansman ortamının hızla geliştirilmesi, sebepleri halen tam olarak anlaşılamayan birçok krize yol açmıştır. 1995 Meksika, 1997 Asya, 1998 Rusya, 2001 Arjantin’de ortaya çıkan krizler, entegre olma çabası içerisindeki gelişmekte olan ülkelere kaynaklanan ve etkilerinin boyutu değişmekle birlikte başka ekonomileri de etkilemiş krizlere örnektir.

Makroekonomik istikrarsızlıklar ve gelişmemiş sığ piyasalar, finansal sistemleri kırılgan duruma getirmekte ve ekonomilerin krize sürüklenmesini kolaylaştırmaktadır. Türkiye’de de küresel piyasalar ile entegrasyon süreci esas olarak 24 Ocak kararları ile başlamıştır. Alınan bir dizi karar ile Türkiye ekonomisi, dış sermaye girişlerine açılmış ve devletin piyasalar üzerindeki etkisi azaltılmıştır. 80’li yıllarda Doğu Blok’unun çöküşü ile başlayan bu süreç günümüzde de devam etmektedir. Türkiye ekonomisi, 1980 yılında dış ticaretin serbestleşmesi ve 1989 yılında uluslararası sermaye hareketleri üzerindeki kontrollerin kaldırılmasıyla finansal entegrasyon sürecinde hızlı adımlar atmıştır. Ancak, ekonomik yapıdaki istikrarsızlıklar nedeniyle, 1990’lı yıllardan itibaren ‘büyüme-istikrarsızlık-kriz’ kısır döngüsüne girmiştir ve bu durum, 2001 krizi ile sonuçlanmıştır. Sonrasında bir dizi karar ile hayata geçirilen ‘Güçlü Ekonomiye Geçiş Programı’ uygulanmaya başlanmış ve mevcut yapısal problemler giderilmeye çalışılmıştır.



Krizlerin ortaya çıkış nedenleri tarihsel süreçte farklılıklar göstermektedir: Örneğin krizler, 18. yüzyıla kadar daha çok savaş, kıtlık gibi olaylara bağlı olarak, 19. yüzyıldan itibaren ise daha çok kapitalist ekonomilere özgü faktörlere bağlı olarak kendini göstermeye başlamıştır. 19. yüzyıl ve sonrası krizlerin, tek bir sektörde ve belli bir bölgede değil, yaygın bir şekilde ve 6-10 yıl gibi düzenli aralıklarla oluştuğu görülmektedir. Artan iletişim teknolojileri ve internet ağları ise her bir krizin etkilerinin, öncekinden çok daha hızlı ve yoğun şekilde hissedilmesi sonucunu doğurmaktadır.

Krizler üzerinde küreselleşmenin etkisi tartışılmaz durumdadır. Özellikle bilgisayar ve bilgi teknolojisindeki yeniliklerin finansal araç ve hizmetlerde yarattığı değişimler ve yeni finansal araç sayısında meydana gelen olağanüstü artış, sermaye akımlarının hızlanmasını kolaylaştırmış fakat aynı zamanda krizlerin yaygınlaşmasında ve etkilerinin daha yoğun hissedilmesinde de hızlandırıcı bir rol üstlenmiştir. Bir ülkede ortaya çıkan krizin diğer ülkelere süratle yayılması küreselleşmenin bir sonucu olarak kabul edilmektedir.

Son yıllarda karşımıza çıkan krizler göz önüne alındığında; (a) artan sermaye hareketliliklerinin, (b) türev araçlar aracılığı ile yapılan çeşitlendirmelerin, (c) riskin genişleyen ve çoğunlukla takip edilemeyen paylaşımının, (d) sistemik risk unsurlarının ve (e) piyasa katılımcıları tarafından verilen hızlı tepkilerin kriz risklerini artırıcı nitelikte olduğunu söylemek mümkündür. Bu risklerin ve olumsuz etkilerin azaltılmasına yönelik çalışmalar uluslararası kuruluşların gündeminde önemli bir yere sahiptir. IMF, WB gibi kurumlar, dönem dönem küresel finansal istikrar raporları yayınlamakta ve gelişmeleri yakından takip etmektedir.

Küresel finans piyasalarına entegre olmuş ülkelerde ortaya çıkabilecek ve tüm sistemi etkileyebilecek problemlerin önüne geçmek için son 40-50 yıllık dönemde ciddi çalışmalar yapılmıştır. İlerleyen bölümlerde, bu çalışmaların temelini oluşturan Basel kriterlerinden ve karşılaşılan zorluklar karşısında alınan tedbirlerden bahsedilecektir.

### 1.1.2. Basel I ve Basel II

Uluslararası düzeyde daha sağlam bir finansal mimari oluşturma çabaları, zorunlu olmamakla birlikte öneri niteliği taşıyan ve yıllar süren çalışmalar ile şekillenen düzenlemeler sayesinde gerçekleştirilmeye çalışılmıştır. Bu kapsamda Basel Komitesi, resmen kurulduğu 1974 yılından beri edindiği tecrübeler ve güncel şartlara bağlı olarak, küresel ölçekte daha istikrarlı bir finansal sistem kurulması amacıyla çalışmaktadır. Basel Komitesi, resmi olarak yasal bir statüye veya otoriteye sahip olmamakla birlikte ilgili ülkelerin kamu kurumlarının üye olduğu bir organizasyondur. Komite tarafından ihdas edilen standart ve ilkeler, büyük ölçüde etkili yönlendirici tavsiyeler niteliğinde olup, dünya genelinde kabul görmektedir. Bu hiç kimsenin piyasaya müdahale etmemesi gerektiğini düşünen, ancak yaşanan olumsuzluklar karşısında da politika üretemeyen mercilerin, resmi olmayan yollar ile piyasaya müdahale etmesi gibi algılanabilir. Komite tarafından günümüze kadar yapılan çalışmaları, beş dönem içinde ele almak mümkündür (Penikas, 2015):

1. 1974 – 1986 – Uzlaş;ı;
2. 1987 – 1998 – Basel I;
3. 1999 – 2008 – Basel II;
4. 2009 – 2011 – Basel III;
5. 2012 – 2014 – Basel III sonrası.

İlk dönemde Komite, mali ve yönetsel açıdan zayıf bankaların sorunlarını çözmek amacıyla çalışan denetçilerin, bankalar ile kurduğu ilişkilere yardımcı olmak görevini üstlenmiştir. Kendi içinde yoğun çalışmalar yapmakla birlikte Komite, bu dönemde hiçbir kamusal duyuru yapmamıştır.

G10<sup>2</sup> ülkelerinden üyelerin oluşturduğu Basel Komitesi'nin finansal sistem üzerindeki etkinliği ilk olarak 1988 Temmuz'unda, esasen bir bankanın batması halinde mevduat sahiplerinin karşılaşılabileceği maliyetleri en aza indirmek için asgari olarak tutulması gereken sermaye üzerine yoğunlaşan bir anlaşma metninin yayınlanmasıyla başlamıştır. Basel I sermaye yeterlilik uzlaşısı, uluslararası arenada faaliyet gösteren bankaların üstlenmiş oldukları riskler ile elde bulundurmaları

---

<sup>2</sup> Belçika, Kanada, Fransa, İtalya, Japonya, Hollanda, Birleşik Krallık, Amerika Birleşik Devletleri, Almanya ve İsveç.

gereken sermaye arasında ilişki kurmaya çalışmıştır. Sermaye Yeterlilik Oranı (SYO), 1988-1992 dönemi için ön görülen tedrici bir geçişle 1992’de asgari %8 olarak uygulanmaya başlanmıştır. Sermaye yeterliğini, beklenmedik olaylar gerçekleştiğinde kayıpları soğuran bir tampon ve doğru şekilde dizayn edilirse, bankaların eylemlerinden doğan risklerin sınırlanmasını sağlayan bir araç olarak görmek mümkündür. Böylece yeni kredi tahsis etmek isteyen banka veya kredi kuruluşları şayet risk katsayılarını tamamlamışlarsa sermaye artırımına gitmek zorunda kalacaklardır. Elbette bu zorunluluğun yeni kredi maliyeti olarak müşteriye yansımaları kaçınılmazdır. Bankaların kaynakları arasında en maliyetlisi sermaye olduğu için de bu artırım, potansiyel olarak finansal sektör içerisindeki rekabete de etki etmektedir (Antão ve Lacerda, 2008).

Uzlaşıda, kredi riskinin yaratacağı sermaye gerekliliği hesaplanırken portföy yaklaşımı benimsenmiştir. Buna göre varlıkların riskleri, dört grup içerisinde değerlendirilmiştir: 1. risk grubu, genellikle İktisadi İşbirliği ve Gelişme Teşkilatı (The Organisation for Economic Co-operation and Development- OECD) hükümetlerinden alacakları kapsamaktadır ve %0 risk ağırlığına sahiptir; 2. risk grubu, genellikle OECD ülkelerinde bulunan bankalardan alacakları içerir ve %20 risk ağırlığına sahiptir; 3. risk grubu, genellikle konut kredilerini kapsamaktadır ve %50 risk ağırlığına sahiptir; ve son olarak 4. risk grubu, genellikle tüketici ve şirketleri kapsamaktadır ve %100 risk ağırlığına sahiptir (Benzin, Trück ve Rachev, 2003). 1993 yılında Komite, tüm G10 ülkelerinde uluslararası bankacılık faaliyeti yürüten tüm bankalarının minimum yeterlilikleri karşıladığını açıklamıştır (BIS, 2017b).

Komite, süreç içerisinde düzenleyici çerçeveyi de yeniden tanımlanmaya başlamıştır. 1995 yılında, bankaların yabancı para, borçlanma araçları, hisse senetleri, mallar ve opsiyonlarından kaynaklanan pazar riskini de sermaye yeterliği hesaplarına eklemişlerdir. Bu dönemde yaşanan bir diğer önemli gelişme de, bankaların piyasa riskinden kaynaklanan sermaye gerekliliklerinin hesaplarında ilk kez içsel modeller (sıkı niteliksel ve niceliksel standartlara bağlı olarak) kullanmalarına izin verilmesidir (Chang, Jiménez-Martin, McAleer ve Pérez-Amaral, 2011).

Basel I’in etkin olduğu dönem, bankacılık başta olmak üzere finans sistemlerinin de-regülasyon ve dünya piyasalarına entegrasyon sürecini kapsamaktadır. Artan rekabet ortamı ve imkanı içerisinde bankalar çalışanlarını

yüksek miktarlarda kredi vermeye zorlamış ve müşterilere ihtiyaçları üzerinde kredi kullandırmaya yönelmiştir (aktaran Dall, 1992, s. 41). Bu dönemde istatistiksel yaklaşımların kullanımında yaşanan hem veri hem de yöntem yetersizliklerinden kaynaklanan başarısız analizler de güvenilir bir kredi ortamının kurulmasında sorunlara yol açmıştır.

Zamanla Basel I, mevcut koşullar ve küresel ölçekte artan piyasa entegrasyonunun getirdiği yeni riskler ile beraber yetersiz kalmaya başlamış ve paralel olarak piyasalarda yaşanan hızlı değişim de bir yeniliğe ihtiyaç olduğunu ortaya koymuştur. Uzlaşya yönelik eleştirilerin odağında, uzlaşının birçok yaklaşımının yeterince sofistike olmayışı ve varlık risklerinin değerlendirilmesinin piyasa tarafından gerçekleştirilmesi ile düzenleyici tarafından gerçekleştirilmesi arasında takılmış olmasıdır (Decamps, Rochet ve Roger, 2004).

Eksiklik ve ihtiyaçlara cevaben yeni metin, 1999 yılında başlayan çalışmalar ve beş yıl süren istişare süreçleri sonucunda 2004 yılında yayımlanmıştır. Basel I, banka sermayesinin toplam miktarına odaklanmıştı ve böylece borcun ödenememesi durumunda ortaya çıkacak risklerin ve mevduat sahipleri üzerindeki maliyetlerin azaltılması hedeflenmişti. Yeni düzenlemeyle ise finansal sistemin güvenliği ve güvenilirliği, bankaların içsel kontrol ve yönetim mekanizmalarına, denetçi inceleme süreçlerine ve piyasa disiplinine daha fazla odaklanılarak sağlanmak istenmiştir (Benzin ve diğerleri, 2003).

Basel II, üç temel yapısal blok üzerine oturtulmuştur. Bunlardan ilki, Basel I tarafından oluşturulmuş standart kuralların genişletilerek yeni bir sermaye yeterliği çerçevesi oluşturulmasıdır. Basel II sermaye yeterlilik kuralları, Basel I'e göre riske daha duyarlıdır. Bu bağlamda sermaye ihtiyaçları, temerrüt olasılığı ve temerrüt halinde kayıp gibi çeşitli risk göstergelerinin tahminlerine bağlı geliştirilen kredi kalitelerine daha bağımlı hale gelmiştir (Antão ve Lacerda, 2008). Risk hesabında getirilen en büyük yenilik, Basel I tarafından tanımlanan kredi ve piyasa risklerinin yanında üçüncü risk unsuru olarak işletme riskinin eklenmesidir.

**Kredi Riski:** Finansal kurumun kredi ilişkisi içerisinde olduğu karşı tarafın sözleşmedeki yükümlülüklerini, sözleşmede yer alan şartlara uygun olarak yerine getirmemesinden kaynaklanan risktir. Kredi riski, karşı taraf riski olarak da

tanımlanabilir. Kredi riski, kurumların nakit akışlarını ve bilançolarını direkt etkilediği için en önemli risk grubu olarak görülmektedir.

Kredi riskini meydana geliş şekli ve kaynağı itibariyle çeşitli alt başlıklar altında değerlendirmek mümkündür (Mermod ve Ceran, 2011):

1. İşlemin sonuçlandırılmama riski: Karşı taraftan kredi geri ödemesinin zamanında ya da tam olarak banka bünyesine alınmaması riskidir. Temerrüt riski olarak da adlandırılmaktadır.
2. Yoğunlaşma riski: Bankaların kredilerinin büyük kısmını belirli bir bölge, sektör ya da firmaya kullandırmasına bağlı olarak ortaya çıkan risktir. Türkiye’de bankanın bir risk grubuna kullandırdığı kredi, banka öz kaynaklarının %10’u veya daha fazlasına erişirse bu kredi büyük kredi sayılır ve Bankacılık Kanunu’na (Kanun numarası: 5411, kabul tarihi: 2005) göre büyük kredilerin toplamı öz kaynak toplamının 8 katını aşamaz.
3. Ülke riski: Kredi müşterisinin başka bir ülkeden olması durumunda, karşı tarafın bulunduğu ülkede yaşanabilecek politik ve ekonomik sorunların, kredi geri ödemelerinde yaratacağı etkileri ifade eder.
4. Maruz kalınan risk: Kredi veren kurumun, yalnızca ana parayı değil kredi işleminden beklenen kârı da içine alan risktir.

**Piyasa Riski:** Bankaların sahip oldukları pozisyonlar itibariyle; finansal piyasalardaki dalgalanmalardan kaynaklanan ve yatırımların beklenildiği kadar kârla ya da en kötü senaryoda zararla sonuçlanmasına sebep olan risk türüdür. Piyasa riski kapsamında kur riski, faiz riski, likidite riski, emtia ve emtialara dayalı türev araç riskleri bulunmaktadır. Piyasa riski sistematik bir yapıya sahip olduğu için kaçınılması zor bir risk türü olarak görülmektedir ancak bankalar, gereken sermaye yeterliliklerini sağlayarak (hesaplanan risklerinin 12,5 katı kadar) faaliyetlerine sağlıklı şekilde devam edebilme şansına sahip olabilirler.

**İşletme Riski (Operasyonel Risk):** Kredi ve piyasa riskleri dışında kalan tüm riskleri içerisinde barındıran risk grubudur. Operasyonel risk kategorileri içinde: (i) insan süreçlerindeki hatalar, (yazılım uygulamalarındaki hatalar, yanlış bilgi içeren raporlar ya da yanlış taraflara yapılan ödemeler), (ii) insan karar hataları (eksik bilgidan dolayı karlı bir ticaretin gereksiz reddedilmesi ya da yanlış ticaret stratejileri

gibi) (iii) yazılım ya da donanım hataları nedeniyle sistem hataları, (veri iletimi ve veri alma sorunları ve yazılım sistem performans hesaplamaları ile eksik veri temelli raporlar üretilmesi), (iv) süreç tasarım hatası (belirsiz bir şekilde tanımlanmış süreç adımlarıyla iş akışı), (v) dolandırıcılık ve hırsızlık (yasadışı faaliyetler ya da kredi kartı sahtekarlığı), (vi) dış zararlar (yangın ve deprem) yer alır. Bu risk çeşidinin ne olduğu ve hesaplanma biçimleri diğer risklere göre daha muğlak olsa da amaç, olabildiğince farklı risk unsurlarını dikkate almaya çalışarak sistemik olarak yaşanabilecek küresel ölçekte sıkıntıların önüne geçmektir. Komite de yeni çerçevenin amacının “risk yönetimine daha fazla önem vermek ve bankaların risk değerlendirme kapasitelerinde devam eden gelişmeleri cesaretlendirmek” olarak belirtmektedir.

Etkin risk değerlendirme ve borç muhasebesi işlemlerinin, oluşturulmuş politikalar ve kurallar dahilinde sistematik bir şekilde yapılması gerekmektedir. Kredi riski derecelendirme sistemleri borç alanın mevcut finansal durumunu ve ödeme kapasitesini dikkate almaktadır ve borç alanın kredi derecesi yeni bir bilgi elde edildiğinde hemen gözden geçirilmeli ve güncellenmelidir (Bank for International Settlements, 2006; prensip 2). Kredi riski değerlendirmeleri ve borç kayıp öngörülerinin ortaya konması için risk ölçme modelleri ve varsayım tabanlı tahminler kullanılmalıdır (Bank for International Settlements, 2006, prensip 3). Yeni uzlaşmada, riskin yeniden tanımlanmış konumu sürekli olarak göze çarpmaktadır.

İkinci yapısal blok, sermaye yeterliklerinin bir denetim otoritesi tarafından incelenmesi ve raporlanmasıdır. Bu inceleme, bankaların sermaye durumunun genel risk profilleri ve stratejileri ile uyumlu olmasını sağlayacak biçimde denetlenmesini (içsel denetim) ve erken müdahale imkanlarının oluşmasını sağlamaktadır. Böylece gerekli durumlarda asgari sermaye yeterliliği oranından daha fazla bir sermaye bulundurma, hesaplamalarda dikkate alınmayan risklerin ortaya konması, stres testleri ile şoklara karşı dayanıklılığın ölçülmesi ve gereken noktalara müdahale edilmesi hususlarında eyleme geçilmesini sağlanmaktadır. Banka denetçilerine, bankaların içsel risk değerlendirme metodlarının uyumu ve sağlamlığını değerlendirmeleri için daha fazla yetki tanınmıştır. Dolayısıyla da denetçiler, risk profilleri dikkat gerektiren bankalar üzerine daha fazla odaklanabilme imkanına sahip olmaktadır. Bir kurumun denetlenme sıklığı, finansal durumuna bağlı olarak daha seyrek ya da sık olabilmektedir.

Basel II'nin son yapısal bloğunu oluşturan piyasa disiplini ise finansal kurumlara, gerekli tüm bilgileri zamanında ve şeffaf bir biçimde açıklama yükümlülüğü getirmektedir. Böylece, birinci ve ikinci yapısal bloklar ile sağlanan istikrarlı ve sağlam yapı, kamuoyunun bilgisine sunularak daha disiplinli ve güvenilir bir ortam oluşması sağlanması hedeflenmiştir.

Basel II, herkese tek beden elbise yaklaşımı yerine, ülkelerin inisiyatiflerine bırakılan ulusal uygulama tercihleri öngörmektedir. Bu itibarla, Basel II uygulamalarının etkinliği, ülkelerin kendi ulusal şartlarına uygun tercihlerini belirleyebilmesiyle sağlanabilecektir.

Basel II, çeşitli nedenlerle eleştirilere de maruz kalmıştır. Örneğin, standart yaklaşımda (bkz. Bölüm 1.2) iyi derecelere sahip kurumların ağırlıkları çok yüksek ve kötü derecelere sahip olanların risk ağırlıkları çok düşük tutulmuştur. Bir diğer eleştiri de içsel denetim yaklaşımlardan birinin (temel veya gelişmiş yaklaşımlar, bkz. Bölüm 1.2) kullanılması için gereken minimum koşulların çok yüksek olmasıdır. Bu da, tüm bankaları kapsamı hedeflenen düzenlemenin uygulanabilirliği ile ilgili kuşku yaratmıştır (Benzin ve diğerleri, 2003). Bir sonraki bölümde, yeni finansal mimari adı altında şekillenen sistem ve buna yönelik güncel eleştiriler sunulmuştur.

### **1.1.3. Yeni Finansal Mimari ve Eleştiriler**

1929 yılında ABD'de başlayan "Büyük Depresyon" dönemi tecrübeleri, özellikle finansal sistemin çekirdeğini oluşturan büyük kurumlar tarafından aşırı risk alma eğiliminin engellenmesini zorunlu kılmıştır. Amaç, benzeri boyutlarda yıkıcı etkileri olan finansal krizlerin mümkün olabilince az ortaya çıkmasını sağlamaktır.

1933 yılında ABD'de yayınlanan ve Glass-Steagall Sistemi olarak da bilinen bankacılık kanunu, ticari bankacılık ile yatırım bankacılığını birbirinden ayırmıştır. Mevduat toplayan ve borç veren bankaların, menkul kıymet işlemleri yapması ya da sigortalaması yasaklanırken, yatırım bankalarının da herhangi bir ticari banka ile yönetim ya da ortaklık gibi birliktelikler oluşturmasının önüne geçilmiştir ("Banking Act of 1933 (Glass-Steagall)", 2013). Kanun aynı zamanda bankaların, ABD Merkez Bankası'na (Federal Reserve System- FED) daha fazla bilgi vermesini ve FED tarafından denetlenmesini de zorunlu kılmıştır. Bu sistem, İkinci Dünya Savaşı'ndan

1970'lere kadar başarılı şekilde uygulanmıştır ve hızlı reel sektör büyümesi ile birlikte azalan işsizlik bu döneme modern kapitalizmin 'Altın Çağı' adı verilmesini sağlamıştır (Crotty, 2009). Tarih boyunca da görüldüğü gibi her şeyin yolunda olduğu dönemlerde geçmiş problemler çabuk unutulmakta, zamanla gevşeyen katılımcılar da yeni krizlerin oluşması için zemin hazırlamaktadır. 1970'lerde, ekonomik iş çevrimi gerçeğine bağlı olarak ABD'de, 60'ların yüksek büyüme ve istihdamının sonucu olarak ortaya çıkan yüksek enflasyon olgusu fiyat istikrarını bozmuş, 1974 ve 1979 yıllarında yaşanan iki büyük petrol şokunun da etkisiyle birlikte kalıcı hale gelmiş ve düşen ekonomik faaliyet ile birlikte işsizlik oranları da yükselişe geçmiştir. Bozulan makroekonomik yapı ve fiyat istikrarsızlığı da finansal kurumların faaliyetlerini gerçekleştirirken ciddi sıkıntılar yaşamalarına sebep olmuştur. Bu dönemde ABD ve diğer büyük kapitalist ekonomiler, finansal piyasaları düzenleyen sistemlerini yeniden gözden geçirerek aşırı risk alma eğilimlerini kısıtlamaya gitme ya da 1930 öncesi dönemin serbest piyasa ideolojisine geri dönme seçenekleri ile karşı karşıya kalmıştır. Etkin finansal düzenlemeleri savunan güçlü bir politik hareket bulunmadığından, İngiliz Başbakanı Thatcher ve ABD Başkanı Reagan önderliğinde, sermaye piyasaları ile ticari bankacılığı ayıran düzenlemeler terk edilmeye başlanmış ve günümüz sermaye piyasası merkezli küresel-boyutta entegre finans sistemi ortaya çıkmıştır. 1970'lerde başlayan ve 1999'da zirvesine ulaşan bu döneme 'Yeni Finansal Mimari' dönemi<sup>3</sup> adı verilmektedir (Crotty, 2009).

'Yeni Finansal Mimari', "Etkin Piyasalar Hipotezi" temel alınarak oluşturulmuştur. Etkin piyasalarda yatırımcıların piyasaya ve dolayısı ile oluşan fiyatlara güveni tam olur. Yatırımcıların, bilginin serbestçe dolaştığı, piyasaya herhangi bir dış müdahalenin bulunmadığı tam rekabet içeren etkin piyasalarda oluşan fiyatlarda anormal (piyasadan farklı) kazançlar elde etmeleri olanaksızdır ve bu sebeple her zaman optimal tercihleri yapacaklardır. Etkin piyasa hipotezine göre fiyatlar her zaman dengededir. Bu görüşle şekillenen yıllar boyunca, finansal piyasalar, özellikle 1990'ların başından sonra hızlanan bir şekilde büyümüştür ancak süreç boyunca bu büyümenin problemleri olduğu, aşırı spekülasyon hareketlerinin piyasalara

---

<sup>3</sup> 'Yeni Finansal Mimari' kavramı, her kriz sonrası oluşturulmaya çalışılan sağlıklı ve istikrarlı yeni bir finans sistemi için kullanılabilir. Bu bölüm yazılırken çoğunlukla Crotty (2009)'dan yararlanılmıştır ve kendisi incelemesinin temelini 2008 öncesi finansal sistem üzerine kurmuştur.



yön verdiđi üzerine çok sayıda işaret de görölmüştür. Dünya Bankası araştırması, 1970'lerin sonu ile 2000'lerin başı arasında dünya çapında 117 sistemik bankacılık krizi tespit etmiştir (Caprio ve Klingebiel, 2003). Bu krizlerin sosyal yansımaları da dikkate değerdir; her şeyin yolunda olduđu zamanlarda kazançlar özel şirketlere/kişilere akarken kriz dönemlerinde oluşan kayıpların maliyeti topluma yansıtılır.

Küresel ölçekte yaşanan finansal krizi diğerlerinden ayıran bazı özellikleri bulunmaktadır. 1929 Buhranı sonrası krizler genellikle gelişmekte olan ekonomilerde ortaya çıkmış, bölgesel nitelikte ve görece sınırlı etki alanına sahip iken küresel finans krizi gelişmiş bir ekonomide başlayarak yayılım göstermiştir. Klasik teori ve yaklaşımlar, krizin sosyal yapılar üzerinde yarattığı yıkıcı etkilerin büyüklüğü, çok kısa bir süre içerisinde birçok dünya ekonomisini etkisi altına alması, kriz ile birlikte küresel ölçekte yaşanan likidite sorunu ve büyük bir kredi çöküşünün yaşanması karşısında yetersiz kalmıştır. Tüm bu sayılan nedenler, 'Yeni Finansal Mimari' yerine geçebilecek bir küresel düzen ve farklı bir finansal mimari arayışı içerisine girilmesini de beraberinde getirmiştir. Finansal sistemin içerisine dahil olmamasına rağmen krizin sebep olduđu olumsuzlukları yoğun şekilde hisseden kitlelerin, hükümetler ve düzenleyiciler tarafından şekillendirilmiş sisteme yönelik yoğun politik eleştirilerinin ortaya çıkması da yeni düzen arayışlarında itici güç olmuştur. Böyle bir ortamda, Basel III konusunda yapılan çalışmalar da bu çabanın bir ürünüdür.

#### **1.1.4. Basel III**

Küresel finans krizi, Basel II uygulamalarının, risk iştahı yüksek ve olabildiğince serbest faaliyet gösteren finansal piyasa katılımcılarının yarattığı sistemik problemleri önlemekte yetersiz kaldığını göstermiştir. Özellikle sistematik öneme sahip kurumların yarattıkları risklerin ve toplumda sebep oldukları maddi/manevi zararların engellenmesinde eksiklikler olduđu ortaya çıkmıştır (Georg, 2011). Mart 2008'de Bear Stearns (1923'te kurulmuş olan ABD merkezli küresel bir yatırım bankası) ile başlayan, ardından Fannie Mae and Freddie Mac'e (ABD konut finansmanına yönelik likidite, istikrar ve ödeme gücü sağlayan devlet destekli kurumlar) sıçrayan ve Eylül 2008'e Lehman Brothers'ın (kuruluşu 1850'lere dayanan ABD merkezli küresel ölçekte bir yatırım firması) iflasını duyurması ile zirveye çıkan

panik; paniğin yarattığı ortamda bankacılık sisteminin çökmesi ve birçok ülkenin, özellikle büyük finans kuruluşlarına ciddi destekler vermek zorunda kalması, sistemin içine düşebileceği kriz durumları için yeterince önlem alınmadığını gözler önüne sermiştir. Finansal sistem ve bankacılık sistemi birkaç sene içerisinde tekrar istikrara kavuşmuş olsa da krizin ölçülebilen ve ölçülemeyen maliyeti çok yüksek olmuştur. Dünya Bankası verisi, 2009 yılında dünya çapında GSYH büyüme oranının (WB, 2017c) -%1.734 olduğunu göstermektedir. Gerekli fon kaynaklarından mahrum kalan reel sektörün de krizden etkilenmesini ile işsizlik oranlarının artmış ve halkların refahında kayıplar yaşanmıştır. Atkinson, Luttrell ve Rosenblum (2013), krizin etkilerini inceledikleri çalışmalarında zenginlik ve refahta yaşanan kayıpların yanı sıra ölçülmesi mümkün olmayan olumsuzluklardan da bahsetmişlerdir. İş kayıpları sebebiyle kişilerde ortaya çıkan stres ve utanç duyguları, finansal yetersizlikler sebebiyle ailelerinden bağımsız bir yaşantı kuramayan genç nüfus sayısında artış, hükümetlere, kurumlara ve kapitalist sisteme duyulan güvenin sarsılması, tam anlamıyla ölçülmesi mümkün olmayan ve iyileştirilmesi ekonomik göstergeler kadar kolay olmayan sonuçlardır.

Basel Komitesi, özellikle sistemik olarak önemli bankaların risklerinin tanımlanması konusunda çalışmalarını FSB ile yürütmektedir. Sistemin kırılabilirliği açısından öneme sahip kuruluşların Basel III çerçevesinin de ötesinde kayıpları karşılayacak kapasiteye sahip olmaları konusunda herkes hemfikirdir. Bankalara, özellikle kısa dönem şoklara karşı dirençlerini artırmaya yönelik likidite karşılama oranı ve uzun vadeli yapısal problemleri önlemek üzere net istikrarlı fonlama oranı tanımlanmıştır. Bu önlemler ile birlikte yüksek maliyetli sermaye kalemlerinin asgari karşılık oranlarının artırılması sebebiyle, banka karlılıklarının ciddi zararlar göreceği yönünde tartışmalar da yaşanmıştır. Önlemlerin ekonomilerde yaratacağı maliyetler konusunda endişeler olsa da (Elliott, 2010) sağlanacak güvenliğin yararları karşısında maliyetlerin düşük kaldığını söylemek mümkündür ve bu durum çeşitli ampirik çalışmalarla da ortaya konmuştur (Allen, Chan, Milne ve Thomas, 2012; Angelini ve diğerleri, 2015; Fidrmuc ve Lind, basımda; Slovik ve Cournède, 2011; Zheng, Rahman, Begum ve Ashraf, 2017).

Basel III olarak adlandırılan düzenleme değişiklikleriyle ulaşılmak istenen hedefler şu şekilde özetlenebilir (Cangürel, Güngör, Sevinç, Kayci ve Atalay, 2010);

- ✓ Kaynağı ne olursa olsun finansal ve ekonomik şoklara karşı bankacılık sisteminin dayanıklılığının artırılması,
- ✓ Kurumsal yönetim ve risk yönetimi uygulamalarının geliştirilmesi,
- ✓ Bankaların şeffaflığının ve kamuya bilgi verme özelliklerinin artırılması,
- ✓ Mikro bazda yapılan düzenlemelerle bireysel olarak bankaların dayanıklılığının artırılması,
- ✓ Makro bazda düzenlemelerle finansal sistemin şoklara karşı direncinin artırılması.

Basel III, Basel II yöntemi yerine getirilmiş ve sıfırdan oluşturulmuş bir düzenleme değildir. Basel II'nin 2007 krizinin oluşumunu önleyememesi, yıkıcı etkilerini soğurmak konusunda eksiklikleri ve bu hatalardan çıkartılan dersler sonucu ortaya çıkmış, devam niteliğinde bir düzenlemedir. Tez çalışmasının temelini oluşturmadığından, Basel III düzenlemesinin ayrıntılarına değinilmeyecektir, daha fazla bilgilendirme için Cangürel ve diğerleri (2010) ve Gürel, Gürel ve Demir (2012)'in çalışmalarına başvurulabilir.

Şimdiye kadar bahsi geçen düzenlemeler kurumların, başta öneri sonrasında ise bir zorunluluk olarak gelişmiş risk yönetim sistemleri kurmaları sağlamıştır. Risk yönetiminin önemli bir parçası olarak risk-modelleme yöntemleri ise, genel itibariyle finansal stresin hiç olmadığı ya da az olduğu dönemlere uygun tasarlandıklarından, kargaşa dönemlerinde başarısız sonuçların alınmasına yol açmışlardır (Rossignolo, Fethi ve Shaban, 2013). Yeni düzenlemelerde temel hedef de risk modellemelerinin iyileştirilmesinden ziyade oluşacak olumsuz durumlarda sistemik riske yol açacak eksikliklerin giderilmesi olmuştur. Finansal sistem içerisindeki katılımcıların risk algularının dönemsel niteliğinin<sup>4</sup>, Basel II ile sunulan içsel denetim yaklaşımları ile ortaya çıkartılmadığı, Basel III ile birlikte söz konusu dönemselliklerin ortaya çıkartılmaya çalışıldığı ve böylece istikrarlı bir sistemin hedeflendiği söylenebilir (Ojo, 2010). Ancak düzenlemenin, kuyruk risklerinin yakalanamadığı içsel denetim

---

<sup>4</sup>Dönemsellik kavramı, ekonomi ve finans literatüründe farklı anlamlarda kullanılabilir. Burada bahsedilen dönemsellik, piyasalarda canlılık olduğu dönemlerde risk algısının düşük, gerileme olduğu dönemlerde ise yüksek olduğu anlamına gelmektedir.

yaklaşımlarının iyileştirilmesinde yeterince katı olmadığı yönünde eleştiriler de mevcuttur<sup>5</sup> (Rossignolo ve diğerleri, 2013).

Şimdiye kadar bahsedilen tüm çalışmalar içerisinde kredi risklerinin doğru şekilde değerlendirilmesi ve buna bağlı olarak kurumların, gereken asgari sermaye yükümlülüklerini yerine getirmesi önemli bir yer tutmaktadır. Bir sonraki bölümde, asgari sermayenin belirlenmesinde önemli bir aşama olan kredi risk değerlendirmesi, kurumlar tarafından yayınlanan raporlara ve yazındaki çalışmalara başvurularak irdelenecektir.

## **1.2. KREDİ RİSK DEĞERLENDİRMESİ**

1980-90'larda bankalar ve kredi derecelendirme kuruluşlarında, kredi derecelendirmesi ve kredi riskinin yönetimi için yeni istatistiksel araçlar kullanılmaya başlanmış ve Basel I ile birlikte yeni bir düzenleyici risk yaklaşımı ortaya konmuştur (Dionne, 2013). Günümüzde çok yüksek düzeylere çıkan veri işleme ve depolama kapasitesi ile daha başarılı risk değerlendirmeleri yapmak mümkün olabilmektedir. Riskli başvurular ile bunların sebep olabilecekleri potansiyel zararları en aza indirmek eskisine nazaran daha kolay hale gelmiştir.

Ancak bu kapasite, kurumların mevcut bilgi ve yaklaşımları zamanında ve yeterli düzeyde uygulamamalarından dolayı tam anlamıyla kullanılamamıştır. Otoriteler Basel süreçleri boyunca doğru risk analizinin önemine vurgu yapmaya devam etmiştir ancak kredilerin değerlendirilmesinde, 90'ların sonuna kadar makine öğrenme yöntemleri ve meta-sezgisel algoritmalar gibi yaklaşımlardan yeterince yararlanılamamıştır. Söylenebilir ki çalışmalar, yeni yöntemlerin ortaya çıktığı ve geliştirildiği tarihleri on yıllara varan gecikmeler ile takip etmektedir. Sorensen, Sevaux ve Glover (2017)'in makalesinde 1990-2000 ve 2004-2012 yılları arası, önemli meta-sezgisel algoritmaların ortaya çıkışının yoğunlaştığı dönemler olarak göze çarpmaktadır. Geriden takip edilen ilk yöntem gelişim sürecinin bugünkü aşamalarını

---

<sup>5</sup> Baştan sona tüm Basel süreçlerinin kapsamlı bir eleştirisi için bkz. Lall (2009).

yakından izlemek gerekmektedir. Tekniklerdeki gelişmeler ile finansal uygulama alanlarının paralel ilerleyişi, kuşkusuz ki büyük avantajlar sağlayacaktır.

Tekniklerin sağladığı faydalar ne kadar iyi olursa olsun, elde edilen sonuçlara göre nesnel karar vermek de bir zorunluluk olarak önemini korumaktadır. Küresel bankacılık sistemi, 2007 senesinde başlayan kriz dönemine, finansal araçlar kanalıyla riskleri tüm sisteme yayılmış değersiz konut kredilerinin yol açtığı problemleri sübvansede edemeyecek düzeyde düşük sermaye yeterlilikleri ile girmiştir. Artan risk iştahı ve yüksek kazanç beklentilerinin sonucunda, doğru ölçülemeyen ve/veya göz ardı edilmiş kredi riskleri ortaya çıkmıştır. Karşı karşıya kalınan durum, IMF, WB, ECB ve Basel Komitesi gibi uluslararası denetleme ve düzenleme kurumlarının finansal istikrarı artırmaya yönelik daha fazla çalışma yürütmesini zorunlu kılmıştır. Bu çalışmalar, iflasların ve likidite risklerinin, sermaye yeterliliklerinin yeniden değerlendirilmesi ile azaltılmasını kapsamaktadır. Sermaye tabanının kalitesini ve seviyesini yükseltmenin yanı sıra, mevcut tüm risklerin sermaye çerçevesi içerisinde yakalanması gerekliliği de ortadadır (BIS, 2010b).

Basel III yapısal blokları, finansal kurumlar için risk yönetimi ve denetiminin geliştirilmesini hedeflemiştir. Bu değişikliklerin, bankalar, kredi birlikleri, varlık yöneticileri ve bireyler gibi piyasa katılımcılarının kârları üzerindeki olumsuz etkileri hakkında endişeler bulunsa da uzun vadede güven ve istikrarın sağlanması çok daha önem arz etmektedir. Değersiz konut kredilerinin yol açtığı karşı taraf temerrütleri sebebiyle bankaların maruz kaldığı riskler, tüketici kredilerinin kalitelerine ve risklerin doğru değerlendirilmesine daha fazla odaklanılması gerekliliğini ortaya koymuştur. Finansal kurumlar arasındaki rekabet, pazar paylarını artırma isteği ve çok büyük veri kümelerinin yeterince hızlı analiz edilememesi gibi etmenler, söz konusu kurumları, doğru şekilde hesaplayamadıkları riskleri kabul etmesine yol açmıştır.

Kredi başvurularının değerlendirilmesi, bahsi geçen sebeplerden dolayı yetersiz analitik incelemelere ve öznel kararlara bağlı olarak gerçekleştirilmiştir. Çok sayıda başvuru çok kısa süreler içerisinde değerlendirilmiş ve yanlış değerlendirmeler, çok büyük boyutlarda yaşanan kayıplara temel oluşturmuştur. Bugün sahip olunan hesaplama yöntem ve kapasiteleri ile yapılabilecek çalışmalar, kişilere bağlı verilen kararlara nazaran büyük avantajlar sağlamaktadır. Bu sebeple karar vericilerin, gelişmiş teknik analizlerden yoğun şekilde yararlanması, öznel yargılarını ve nihai

kararlarını güçlendirmeleri gerekmektedir. Uluslararası kurumlar tarafından geliştirilen ve düzenleyici nitelikteki önerilerin yenilikçi tarafı, karar vermenin temeline bireysel tecrübe ve kıdemli yönetici fikirleri yerine istatistiği yerleştirmesidir. Böylece, kaliteli veriye sahip olmak ve analiz gücü gerekliliklerini karşılamak, risk yönetiminin ana akımı haline gelmeye başlamıştır.

Claessens ve Kodres (2014) tarafından da belirtildiği gibi, sistemik riskleri azaltmak için bir sistem yaklaşımı geliştirmek gerekmektedir. Varlıklara bağlı risklerin ve borç verenlerin durumlarının takibi bu sistemik yaklaşımın bir unsurudur. IMF raporlarında, banka sermayelerinin miktar ve kalitesine yönelik makro-ihtiyati politikaların yeni çağına işaret edilmektedir (Fund, 2013a, 2013b). Zayıf banka sermayeleri krizlerin içerisinde domino etkilerinin en önemli sebeplerindedir. Claessens ve Kodres (2014), yüksek riskli, likit olmayan ve uzun vadeli varlıklara sahip bankaların risk alma eğilimleri sebebiyle, önceki dönemlere göre daha fazla sermaye yeterliliklerine ihtiyaç duyulduğunu belirtmişlerdir.

Finansal kurumların maruz kaldığı risklerin değerlendirilmesi ve yönetilmesi, düzenleyiciler tarafından standartlaştırılması gereken konulardır (BIS, 2000). Kredi riskinin değerlendirilmesi için uluslararası düzeyde kabul görmüş üç seçenek bulunmaktadır: standart yaklaşım, temel yaklaşım ve gelişmiş yaklaşım (Ferguson, 2003).

*Standart yaklaşım:* Bu yaklaşım temel olarak, bankalardaki kredi riskinin bağımsız kredi derecelendirmelerine dayanılarak standart bir biçimde ölçülmesidir (Cangürel, 2012, s:55). Standart yaklaşım varlıkları, denetimci şartlarına bağlı olarak sabit risk ağırlıkları ile kategorize etmekte ve bu kategoriler tüm bankalar arasında aynı şekilde kullanılmaktadır. Örneğin başka bankalara verilen kredilerde, AAA kredi notuna sahip kurum için uzun vadede %20 (kısa vadede %20), BBB kredi notuna sahip bir kurum için ise uzun vadede %50 (kısa vadede %20) risk ağırlığı uygulanmaktadır. Kredi verilen kurumun türüne ve kredi notlarına bağlı olarak risk ağırlıkları %10'lardan %150'lere kadar değişkenlik gösterebilmektedir.

Bu yöntem gelişmiş modelleme çalışması gerektirmediğinden ve basit olduğundan özellikle daha az deneyime sahip küçük bankalar tarafından tercih edilmektedir. Basitliğine karşın standart yaklaşım, esnekliğinin ve gerçekçiliğinin az ve çoğunlukla çıkar çatışması durumları içerisinde görev yapan kredi derecelendirme

kuruluşlarının dışsal analizlerine bağlı olması nedeniyle eleştiri konusu olabilmektedir.

*İçsel değerlendirmeye dayalı yöntemler:* İçsel değerlendirmeye dayalı yöntemler temel ve gelişmiş yaklaşım olarak ikiye ayrılmaktadır. Temel yaklaşımda kurum, borcun temerrüt olasılığını belirler ve geriye kalan değerlendirme denetimci tarafından yapılır. Gelişmiş yaklaşımda ise tüm kredi değerlendirme sürecinin sorumluluğu kuruma verilir. Bu yaklaşımlar, otoritelerin onayı ve kontrolü ile kurumların kendi kredi risklerini, oluşturdukları içsel modeller ile ölçerek sermaye gereksinimlerini hesaplamalarına dayanmaktadır. Özellikle kredi portföyünde konut kredileri ve yatırım yapılabilir ticari kredi oranı yüksek olan bankaların, Basel I ile karşılaştırıldığında, sermaye gereksiniminde %30'lara yaklaşan düşüş sağlayabileceği ileri sürülmektedir gerekmektedir (Tattersall ve Smith (2005)'ten aktaran, Cangürel, 2012).

Bankalar içsel değerlendirme ile çeşitli risk parametrelerini hesaplayarak risk ağırlıklarını belirler. En önemli risk parametreleri temerrüt olasılığı, temerrüt halinde kayıp ve temerrüt tutarıdır. Temerrüt olasılığı, kredinin bir yıllık zaman dilimi içerisinde ödenmemesine ilişkin tahmin parametresidir ve 0-1 değerlerinden birisini almaktadır. Temerrüt halinde kayıp, kredinin temerrüt olasılığı 1 iken kurumun maruz kalacağı kayıp tutarının tahmin edilmesidir. Temerrüt tutarı, kredi alanın temerrüde düşmesi halinde kredinin beklenen risk tutarı olarak tanımlanmıştır. Tüm bu parametreler belirlendikten sonra kurum, bu tahmin edilen bu üç değeri birbiriyle çarparak beklenen kayıp tutarına ulaşır. Kurumdan beklenen, tüm bu beklenen kayıp tutarını karşılayacak kadar sermayeye sahip olmasıdır.

İçsel değerlendirme yöntemleri çok daha gerçekçi varsayım ve analizlere dayanmaktadır. Büyük bankalar tarafından tercih edilmekte ve ABD'de kritik kuruluşlar için zorunlu tutulmaktadır. Ancak daha önce de bahsedildiği gibi bu modeller, nadir olayların yaratacağı riskleri yakalamakta başarısız olmakta, gereken sistemlerin kurulması ve sürekliliği için büyük yatırımlar ve analiz için çok sayıda veri gerekmektedir.

İçsel denetim yaklaşımı ile değerlendirme yapan kurumlar için kredilerin derecelendirilmesi ve değerlendirilmesine yönelik sürekli artan başarılarla sahip yöntemler bulunmaktadır. Modern teknolojiyle birlikte yapay zekâ, numerik

matematik, istatistik ve sezgisel yaklaşımlarda meydana gelen gelişmeler, risk tahminlerinde umut vaat eden çalışmaların ortaya çıkmasına da vesile olmuştur. Bu yaklaşımlar, az sayıda veri ile de olumlu sonuçlar verebilmektedir. Sonraki iki bölümde, bu yöntemlere ilişkin yazın taraması sunulmuştur.

### **1.3. İSTATİSTİKSEL VE MAKİNE ÖĞRENME YÖNTEMLERİ**

İnsanlığın bilgi üretimi ve üretilen bilgiyi saklama kabiliyeti son yarım yüzyıllık süreç içerisinde muazzam bir artış göstermiştir. Bilişim teknolojileriyle paralel olarak da veriyi işleyen ve işlenen veriden anlam çıkartarak karar vericilere destek olan yöntem ve araçlar da gelişmeyi sürdürmüştür. Kullanıcılar bu yöntem ve araçları, matematik ve algoritma detayları ile uğraşmak zorunda kalmadan paket programlar desteğiyle rahatlıkla uygulayabilmektedir. Bu kapsamda istatistiksel ve makine öğrenme teknikleri de birçok alanda başarıyla kullanılmaktadır. İstatistiksel öğrenme yöntemleri çeşitli varsayımlar altında çalışırlar ve hipotezler kurarlar. Örneğin doğrusal regresyon (DR), bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki olduğu, hata değerlerinin normal dağılım özellikleri gösterdiğini ve regresyon doğrusu çevresindeki değişkenliğin tüm tahmin değişkenleri boyunca aynı olduğu varsayımı altında çalışır. Modelin oluşturulması sırasında da matematiksel eşitlikler kullanırlar. Makine öğrenme yöntemleri ise öğrenmelerini bu tür varsayımlardan bağımsız olarak veriyi çeşitli yönergeler dahilinde analiz ederek gerçekleştirirler. Örneğin RO veriye ilişkin hiçbir varsayımda bulunmadan ayrıştırma adım ve kurallarını takip ederek çalıştırılır. Tezin bundan sonraki kısımlarında kısaca öğrenen yöntemler olarak adlandırılacak olan bu tekniklerin ortak noktası veri üzerinde öğrenme gerçekleştirerek, bağımlı değişken ile bağımsız değişkenler arasındaki örüntüleri ve bağlantıları ortaya çıkartmaya çalışmalarıdır. Bağımlı değişkenin olmadığı durumlarda öğrenme, tüm değişkenler arasındaki ilişkilerin analizi ile gerçekleştirilir. Çıktıların yorumlanmasına bağlı olarak üretilen bilgi, analizin amacına uygun olarak değerlendirilmektedir. Amaç, yalnızca öğretilen veri kümesi üzerinde daha iyi bilgi sahibi olmak ve değişkenler arası ilişkileri anlamak olabileceği gibi, öğrenilen bilgiye bağlı olarak tahmin yapmak da olabilir. Liao, Chu ve Hsiao (2012), daha çok anket ve klasik istatistiksel yöntemlerin tercih edildiği sosyal bilimler



alanlarındaki çalışmalarda veri madenciliği yöntemlerinin (öğrenen yöntemlerin) kullanışlı olabileceğine işaret etmektedir.

Öğrenilen veri kümesinin daha iyi anlaşılması ve gözlemler/değişkenler arası ilişkilerin ortaya çıkartılmasına öğrenme terminolojisinde *yorumlama*; öğrenme ile kazanılan genelleme yeteneği ile, farklı veri kümelerinde aynı değişkenlere bağlı olarak değer alan bağımsız değişkenin hesaplanmasına ise *tahminleme* adı verilmektedir. Öğrenen yöntemler, yorumlama ve tahminleme başarıları ölçülerinde değişen performanslara sahip olabilmektedir. Örneğin, klasikleşmiş bir yöntem olarak DR, öğretilen veri kümesindeki ilişkileri yorumlama çalışmalarında başarılı sonuçlar verebilmekle birlikte, yeni bir veri kümesi ile tahminleme çalışmalarında o kadar başarılı olamamaktadır. İyi bir tahmin yapmak, daha *esnek* öğrenme mekanizmalarına sahip yöntemler ile mümkün olabilmektedir. Bu yöntemlere örnek olarak da yine birçok alanda uygulaması bulunan destek vektör makineleri (DVM) gösterilebilir. Kabaca söylenebilir ki öğrenen yöntemlerin yorumlama ve tahminleme yetenekleri arasında bir değiş-tokuş varlığı söz konusudur ve analistler, amaçlarına bağlı olarak uygun model(ler)i belirlemeli ve kullanmalıdır.

Ele alınana probleme göre tercih edilen öğrenen yöntemleri, eğitim sırasında bir *bağımlı değişken* kullanımı olup olmadığına göre *denetimli ve denetimsiz* olmak üzere iki gruba ayırmak mümkündür. Bu kategoriler, bir öğretmen desteği ile öğrenmek ya da destek almadan öğrenmek olarak da tanımlanabilir.

Denetimli öğrenimde “öğretmen” hedef değeri sisteme sunar ve sistem bu hedef ile kendi çıktısı arasındaki hataya bağlı olarak kendini düzenler (Altınbas ve Biskin, 2015). Denetimli öğrenimde, *tahmin (bağımsız) değişkenleri* kullanılarak bağımlı değişkeni en iyi açıklayan bir model oluşturulmaya çalışılmaktadır. En iyi açıklama ölçütü ise, bağımlı değişkenin sayısal bir değere sahip olması durumunda tahmin edilen değerler ile gerçek değerler arasındaki farkın en küçük olması; bağımlı değişkenin kategorik bir sınıflandırma değerine sahip olması durumunda ise hatalı tahmin sınıflandırmalarının en küçük olmasıdır.

Denetimsiz öğrenimde ise tahmin edilmeye çalışılan bir hedef değişken yoktur. Bunun yerine yöntem, veri kümesindeki değişkenler arasında analist tarafından modellenmemiş ve gizli ilişkileri inceleyerek gözlemleri sınıflandırmaktadır. Elde edilen bilgi de kullanım amacına göre değerlendirilir: Örneğin ekonometrik bir

analizde, makroekonomik deęişkenler arasındaki iliřkilerin mevcut teoriler ile uyumu veya bir pazarlama alıřmasında tüketicilerde harcama tercihleri arasında teorilere baęlı olmayan yeni baęlantıların ortaya ıkartılması gibi.

Bu alıřmada denetimli öğrenme dizaynına uygun bir tüketici kredi veri kümesinin analizi amaçlandıęı için yalnızca kredi derecelendirme ve deęerlendirme alıřmalarında kullanılan denetimli öğrenme yöntemleri yazınına yer verilmiřtir.

Kredilerin kategorik olarak sınıflandırılmasında, lojistik regresyon (LR), doğrusal ayrımcılık analizi (DAA), k-en yakın komřu sınıflandırıcısı (KEK), RO, hızlandırılmıř ağalar (HA) ve DVM akademik ve pratik alıřmalarda sıklıkla kullanılan yöntemler arasında yer almaktadır. Bu yöntemlerin yanı sıra; doğrusal, karesel, ok kriterli, tam sayılı, dinamik programlamalar ve makine öğrenme yöntemleri de uygulamalarda kendilerine yer bulmaktadır. Yazında, bu yöntemler arasından bir “en iyi” seimi yapabilmek mümkün deęildir; her yöntem, ele alınan probleme ve veri kümesine baęlı olarak avantajlara ve dezavantajlara sahiptir ve kurulan modeller ile veri analiz ve sonuçlar test edilmeden bir yöntemin dięerine göre üstünlüğünü anlamak mümkün deęildir.

Sınıflandırma alıřmalarında kullanılan öğrenen yöntemler parametrik ve parametrik olmayan yöntemler olarak gruplandırılmaktadır. İzleyen bölümlerde, kredi risk deęerlendirme alıřmalarındaki uygulamalara iliřkin yazın sunulmuřtur.

### **1.3.1. Parametrik Yöntemler**

Parametrik yöntemler, tahmin deęişkenlerinin bir fonksiyonu olarak ifade edilir ve baęımlı deęişken hakkında sağladıkları sistematik bilginin formu hakkında varsayımlara dayalı olarak alıřırlar. Örneęin, DR’da, söz konusu iliřkinin doğrusal olduęu varsayılır ve tahmin problemi, tüm deęişken boyutlarının incelenmesi için ayrı ayrı fonksiyonların oluřturulması yerine deęişken katsayılarının tahmin edildięi tek bir boyuta indirgenir. Finans alanında yapılan sınıflandırma alıřmalarında kullanılan yöntemlerin ilk örnekleri arasında parametrik istatistiksel yöntemler yer almaktadırlar.

Örneęin Altman (1968), dönemin řirket iflaslarının tahmininde kullanılan geleneksel oran analizi performansını, girdi olarak finansal ve ekonomik oranları kullandıęı oklu ayrımcılık analizi (AA) ile karřılařtırmıřtır. Yalnızca üretim

firmalarını dahil ettiği çalışmada, geleneksel ve basit yöntemlerin yeterli olmadığını, ÇAA ile %94-95'lere varan oranlarda tahmin başarısı elde ettiğini belirtmiştir. Çalışmasının kısıtı olarak ise ele aldığı firmaların hepsinin kapsamlı finansal verilerinin olduğunu ve bunun diğer sektörlerde faaliyet gösteren firmalarda aynı ölçüde geçerli olamayabileceğini bildirmiştir. Hangi alanda tahminleme çalışması yapılırsa yapılsın, kapsamlı ve doğru verinin varlığının ve niceliğinin önemini sürdürdüğü günümüzde de söz konusu kısıt geçerliliğini korumaktadır.

Çalışmalarında kredi verilerinin doğasına ilişkin yorumda bulunan Reichert, Cho ve Wagner (1983), ÇAA'nin kredi derecelendirme çalışmalarında kullanımına yönelik yaptıkları yeterlilik çalışmasında, kredi veri kümelerinin incelendiği model çıktılarının genelde varsayımlara karşı duyarlılık göstermediğini ve ÇAA tahmin performansının yetersiz olduğunu, kredi sınıflandırma problemlerinde ÇAA yerine DR, faktör analizi gibi başka istatistiksel teknikler kullanmanın daha iyi sonuçlar verebileceğini belirtmişlerdir. Reichert ve diğerleri (1983), analizin tam olarak sonuca ulaşabilmesi için insan deneyim ve yargısının da önemli olduğunu vurgulamaktadır. ÇAA performansının ancak belirli varsayımlar altında başarılı olduğu, Karels ve Prakash (1987) tarafından da onaylanmıştır. Örneğin açıklayıcı değişkenlerin ortak çok değişkenli normal dağılıma uygun hareket etmeleri, sınıflandırma gruplarının benzer dağılım matrislerine sahip olmaları ve analiz öncesi sınıf olasılık ve yanlış sınıflandırma cezalarının doğru şekilde belirlenmesi gerekmektedir.

Desai, Crook ve Overstreet (1996), sinir ağları (SA), DAA ve LR kullanarak Güneydoğu Amerika'da faaliyet gösteren üç kredi birliğinden topladıkları 2733 bireysel kredi başvurusunu değerlendirmiştir. Performans ölçütünün "iyi" ve "kötü" (bkz. Bölüm 2.3; söz konusu çalışmada "kötü" etiketi son 48 ay içerisinde hesabı silinen ya da batan müşterilere verilmektedir.) müşterilerin hepsinin başarılı şekilde sınıflandırılması olması durumunda LR yöntemi ile başarılı sonuçlar elde edilebildiğini belirtmişlerdir. Gayri resmi kredi risk değerlendirme yöntemlerinin yerine kredi derecelendirme yöntemlerini öneren Dinh ve Kleimeier (2007), Vietnam bankalarının bireysel kredi riski yönetimlerinde LR gibi yöntemlere başvurmalarının hem kârlılıklarında hem de rekabette büyük avantajlar sağlayabileceğini söylemişlerdir.

Parametrik yöntemler, özellikle 2000’li yılların başından itibaren yaygın kullanım özelliğini kaybetmiştir. Bunda, başta SA olmak üzere yüksek işlem gücü gerektiren yöntemlerin popülerlik kazanması büyük rol oynamıştır. N. Chen, Ribeiro ve Chen (2016)’nın kredi risk değerlendirme çalışmalarına ilişkin yaptıkları tarama çalışmasında da DVM, SA ve hibrit tekniklerin son yıllarda artan başarılı kullanım örneklerine dikkat çekilmektedir. Buna rağmen parametrik yöntemler, özellikle karşılaştırma amaçlı çalışmalarda halen önemli yer tutmaktadır ve zaman zaman yeni önerilen yöntem ya da yöntemlerden çok daha iyi sonuç verebilmektedir.

### 1.3.2. Parametrik Olmayan Yöntemler

Parametrik yöntemler, dizaynları ve testleri gereği popülasyon hakkında varsayımlarda bulunurlar. Bu varsayımlar doğruya yakın olsa bile, pratikte, incelenen veri kümesinin temsil ettiği ana kütledeki ilişkinin formu ve şekli hakkında kesin olarak bilgi sahibi olmak neredeyse imkansızdır. Kesin bilgi sahibi olamadığımız ana kütle üzerinde varsayımlarda bulunarak eldeki verinin sınanması, doğru olmayan sonuçların elde edilmesine ya da hiçbir sonuç elde edilememesine sebep olabilir. Birçok sosyal bilim dalında, örneklem üzerinden ana kütle hakkında genelleme yapmak bir ihtiyaç olabilmekle birlikte, %100 kesinliğe ulaşmanın neredeyse imkânsız olduğunu unutmamak gerekir. Mevcut örneklemin kısıtlamalara bağlı olmadan incelenmesi ve yorumlamaların da yalnızca bu örneklem üzerinde yürütülmesi çoğu zaman daha sağlıklı sonuçlar verecektir. Bu anlamda parametrik olmayan yöntemler ile örneklemin, ana kütle üzerinde varsayım ve parametreleri karşılamaya gerek kalmaksızın analizi gerçekleştirilir. Parametrik olmayan yöntemlerin en önemli dezavantajı ise değişkenler arası ilişkilerin güvenilir düzeyde tahmin edilebilmesi için gereken gözlem sayısının, parametrik yöntemlere göre çok daha fazla olmasıdır.

Sınıflandırma çalışmalarının birçoğu ve özellikle son yıllarda yapılanlar, parametrik olmayan yöntemlerin parametrik yöntemlere göre daha iyi performans sergilediğini göstermektedir. Desai ve diğerleri (1996), Bölüm 1.3.1’de bahsi geçen çalışmalarında SA’nın, yalnızca “kötü” kredi başvurularının sınıflandırılma performansı ölçüt alındığında diğer yöntemlere göre çok daha iyi sonuç verdiğini belirtmişlerdir. Bunun sebebi olarak, gözlenme oranı genellikle daha çok olan “iyi”

kredilerin öğrenme sürecinde ağır basması ve sistemi yanlılığa düşürmesi gösterilebilir. Temerrüde düşen ve düşmeyen kredi kullanıcılarının veri kümesindeki dengesizliği, öğrenen yöntemlerin tahmin performanslarını etkilemektedir (Bu çalışmada da benzer bir durum söz konusudur, bkz. Bölüm 0). Bu şekilde dengesiz dağılıma sahip veri kümeleri üzerinde yaptıkları analizlerde Brown ve Mues (2012), RO ve gradyan güçlendirme sınıflandırması yöntemlerinin başarılı sonuçlar verdiğini bulmuşlardır. Dengesizliğin daha fazla olduğu durumlarda ise karar ağaçlarının (KA), karesel ayrıştırma analizinin (KAA) ve KEK yöntemlerinin kötü performans gösterdikleri görülmüştür. RO yönteminin LR, KEK ve torbalanmış KEK yöntemleri karşısında daha başarılı olduğunu söyleyen bir diğer çalışma da Kruppa, Schwarz, Arminge ve Ziegler (2013) tarafından yapılmıştır. Yüz binin üzerinde gözlem ve 181 farklı değişken içeren bir veri kümesinde kredi risk değerlendirmesi yapan Addo, Guegan ve Hassani, (2018), LR, RO, gradyan artırma ve derin öğrenme (SA yöntemlerinden oluşan gelişmiş bir öğrenme süreci) yöntemlerini kullanmıştır. Analiz sonuçları karşılaştırıldığında RO yönteminin en iyi olmasa da en iyiye yakın sonuçlar verdiği görülmüştür. Flood (2017) de, kredi kartı başvurularını değerlendirdiği çalışmasında RO ile yüksek tahmin başarısı elde etmeyi başarmıştır. Van Sang, Nam ve Nhan (2016), kredi değerlendirme için paralel çalışacak şekilde dizayn edilmiş RO yöntemi kullanmıştır. Yazarlar yöntemlerinin, RO tek başına kullanımı ve DVM, KA, KEK, Bayes ve SA yöntemlerine göre daha iyi sonuç verdiğini bildirmişlerdir.

SA, insan beyninin belirli bir işi ya da fonksiyonu yerine getirme işlevinden esinlenerek geliştirilmiştir (Haykin, 1999). Bir SA, beynin bilgi işlemesi gibi biyolojik sinir sistemlerinin bilgi işlem süreçlerini taklit eder. SA modellemesinde, beyindeki nöronların işleyişine benzer şekilde işlev gören çok sayıda nöron tanımlanmakta ve sisteme yeni giren bilgiyi birbirlerine aktarıp kendilerini düzenleyerek öğrenme sürecini gerçekleştirmektedirler.

2000'ler, kredi sınıflandırma yazınında özellikle SA'nın sıklıkla kullanıldığı bir dönem olmuştur. Örneğin, klasik bankacılık ve kredi sistemine erişim zorluğu yaşayan birey ve küçük işletmelere yönelik hizmet sağlayan kurumlarını kapsayan 5500 kişilik veri setinde SA ile DAA, KAA ve LR yöntemlerinin performanslarını karşılaştıran Blanco, Pino-Mejías, Lara ve Rayo (2013), SA'nın diğer yöntemlere göre çok daha iyi sonuçlar verdiğini söylemiştir. Aynı veri kümesi üzerinde yapılan başka

bir çalışmada da (Cubiles-De-La-Vega, Blanco-Oliver, Pino-Mejías ve Lara-Rubio, 2013) SA'nın, çok başarılı sonuçlar verdiği ve şirket maliyetlerini ciddi oranda düşürme potansiyeli olduğu belirtilmiştir.

Geleneksel yöntemler (DAA ve LR) ile SA'nın ve GA'nın karşılaştırıldığı bir çalışmada (Desai, Conway, Crook ve Overstreet Jr., 1997), geleneksel yöntemlere göre SA ve GA performanslarının istatistiksel olarak daha iyi olmadığı söylenmiştir. Yazarlar bu sonucun, analiz edilen veri kümesinde doğrusal olmayan ilişkilere sahip değişken varlığının az olmasından kaynaklanabileceğini de not düşmüşlerdir. Geleneksel yöntemler ile çeşitli SA mimarilerinin (nöron kurulum ve bilgi aktarım şemaları) karşılaştırıldığı bir çalışmada da (West, 2000) benzer şekilde, özellikle LR'un kredi başvurularının değerlendirilmesinde başarılı tahmin sonuçları elde edilmesini sağladığı, uzmanlık gerektiren SA modellemelerinin ise çok fazla iyileştirme sağlamadığını belirtilmiştir. Önceden de belirtildiği gibi sınıflandırma amaçlı bir veri kümesi incelemesinde en başarılı yöntemin belirlenebilmesi için, bu tip sınıflandırma çalışmalarında kullanılacak farklı yöntemlerin denenmesi ya da benzer veri kümeleri üzerinde yapılan çalışmaların sonuçlarının karşılaştırılması gerekmektedir.

SA, genellikle uzun eğitim süreleri ve yüksek uzmanlık gerektirecek düzeyde bilgiye bağlı olarak belirlenen ağ topolojilerinin doğru seçimine olan duyarlılığı gibi sebeplerle eleştiri konusu olabilmektedir. Lee, Chiu, Lu ve Chen (2002) SA ve ayrıştırma analizi ile hibritlenmiş bir yöntem önererek hem işlem süresini azaltmayı başarmış hem de ayrıştırma analizi ve LR gibi geleneksel yöntemlere göre daha iyi sonuç elde etmeyi başarmışlardır. Bu tekniklerin tek başlarına başarı elde edilmesinde etkili olmadığını, gelişmiş modelleme bilgi ve becerisinin de sürecin önemli bir parçası olduğunu göstermektedir. Örneğin, analizlerinde çok değişkenli uyarlanabilir regresyon kesitleri ile hibritleştirilmiş SA kullanan T.-S. Lee ve Chen (2005), yöntemlerin tek tek kullanımına veya ayrıştırma analizi ve LR'a göre çok daha iyi sonuçlar elde etmeyi başarmışlardır.

SA, parametrik olmayan teknikler arasında sınıflandırma için yüksek performans sağlayan yöntemler olarak görülmektedir. Ancak başvuranın neden ret edildiği bilgisinin olmayışı, SA'nın kullanımı ile elde edilen sonuçlara göre hareket eden karar vericilerin karşılaştığı önemli bir problemdir. SA, kara kutu adı verilen ve

yalnızca girdinin ve çıktının bilinebildiği yöntemler arasında yer almaktadır. Yani, bir başvurana başvurusunun ne sebeple ret edildiğine ilişkin bilgi vermek amacıyla yorumlanabilir bir bilgi sunmamaktadırlar. Bu eksikliğe karşı Baesens, Setiono, Mues ve Vanthienen (2003), hem SA'nın performanslarından yararlanmak hem de yorumlanabilir bilgi elde etmek amacıyla çalışmalarında çeşitli kural çıkarma teknikleri kullanmışlardır. Kural çıkarma teknikleri, kara kutu adı verilen ve öğrenme süreçlerinin kullanıcıya sunulmadığı yöntemlerden yararlı bilgi çıkarmaya yarayan araçlardır. SA'nda da kural çıkarma teknikleri, kara kutu sistemi olan SA'nı, içsel bilgiyi sembolik kurallar kümesine çevirerek beyaz kutu sistemine çevirir (Taylor ve Darrah, 2005). Böylece karar verici, başvurunun ret sebeplerine yönelik bilgi sahibi olabilecektir. Kredi başvurusunda bulunan müşterileriyle birebir ilişki halinde olan finans sektöründe bu tür bilgiye ulaşmak oldukça önemlidir. Hangi yöntem kullanılırsa kullanılsın, yorumlanabilir bilgi elde etmenin de bir amaç olarak ele alınması gerekmektedir.

DVM kullanımı da son yıllarda SA kullanımı kadar popüler hale gelmiştir. Bellotti ve Crook (2009), DVM ile, LR, DAA ve KEK gibi yöntemlere göre daha başarılı sonuçlar elde etmeyi başarmışlardır. DVM yönteminin yazına yeni yeni girmeye başladığı bir dönemde Z. Huang, Chen, Hsu, Chen ve Wu (2004), Amerika ve Tayvan ekonomilerinde faaliyet gösteren firmaların kredi skorlarının doğru değerlendirilmesi üzerine yaptıkları çalışmada DVM'nin SA'lar ile karşılaştırılabilir sonuçlar verdiğini göstermişlerdir. Çalışmada, ülke ekonomileri arasında değerlendirme için kullanılan değişkenlerde benzerlikler olduğunu ancak bu değişkenlerin önemlerinde farklılaşmalar olduğuna işaret edilmektedir. Bu sonuç, veri kümelerinde sınıflandırma çalışması yapılan tüm analizlerde geçerli olan bir noktayı vurgulamaktadır: Analiz sonuçları eldeki örnekleme göre değerlendirilmeli ve genelleme yapmaktan kaçınılmalıdır. Aynı ayrı veri kümeleri analiz edilmeden yapılacak her yorum veya alınacak her kararın yanıltıcı olma ihtimali bulunmaktadır. DVM ile yüksek performans elde edilen bir diğer çalışma da Harris (2015) tarafından gerçekleştirilmiştir.

Kredi değerlendirme çalışmalarına bakıldığında RO yönteminin, eski çalışmalarda kendisine çok yer bulamamakla birlikte son yıllarda artış gösteren popülariteye sahip olduğu görülmektedir. Bir sonraki bölümde, özellikle büyük veri

kümelerinin analiz edilmesinde yüksek fayda sağlayan RO yönteminin başka alanlardaki konular da dahil olmak üzere finans alanındaki kullanımlarına ilişkin bilgiler verilmiştir.

### 1.3.3. Rastgele Ormanlar Yönteminin Tercih Edildiği Çalışma Konuları

RO yöntemi, gözlem sayısının az ve/veya gözlemler ile ilgili toplanan özellik sayısının çok fazla olduğu tıp, biyoloji, biyo-informatik ve görüntü işleme gibi alanlardaki çalışmalarda kendisine sıklıkla yer bulmaktadır. Tablo 1’de, RO kullanımına ilişkin yazında dikkat çeken çalışmalar sunulmuştur.

**Tablo 1:** Rastgele ormanlar yönteminin kullanım konu ve amaçları

Çalışma/kaynak	Çalışma konusu	Amaç
Sylvester ve diğerleri (2018)	Ticari amaçlı yetiştirilen somon balıkları	Soy yapılarının tespiti
Altınbas, Pacelli ve Sica (2018)	Avro bölgesi devlet tahvil getirileri	Getirileri etkileyen makro-ekonomik değişkenlerin belirlenmesi
Basak, Kar, Saha, Khaidem ve Dey (basımda)	Hisse senedi piyasaları	Endeks fiyat tahmini
Valletta, Torney, Kings, Thornton ve Madden (2017)	Ulusal bir parkta bulunan Afrika antilopları	Görüntüler yardımıyla nüfus takibi
Voyant ve diğerleri (2017) - tarama	Yenilenebilir enerji	Güneş ışınımı tahmini
Sarica, Cerasa ve Quattrone (2017) -tarama	Alzheimer hastalığı	Görüntüler yardımıyla hastalığın erken teşhis ve takibi
Cavalcante, Brasileiro, Souza, Nobrega ve Oliveira (2016) -tarama	Finansal Piyasalar	Tahmin ve karar verme
Liu, Chan, Alam Kazmi ve Fu (2015)	Mali dolandırıcılık	Dolandırıcılık yapan şirketlerin tespit edilmesi
Kumar ve Thenmozhi (2014)	Hisse senedi piyasaları	Endeks fiyat tahmini
Kruppa, Ziegler ve König (2012) - tarama	Türler arası genom ilişkileri	İlişkilerin ortaya çıkartılması sonrası hastalıklara bağlı risk tahminleri
De Luca, Riviaccio ve Zuccolotto (2010)	Portföy seçimi	Kriz döneminde finansal varlık getirilerinin hesaplanması



Bu tür alanlarda incelenen veri kümelerinde bağımlı değişken ile tahmin değişkenleri arasındaki ilişki çoğunlukla karmaşık yapıdadır ve tahmin değişkenleri kendi aralarında yüksek ilişki düzeylerine sahiptir ve bu durum RO gibi yöntemlerin tercih edilmesini (ayrıntılı bilgi için bkz. Bölüm 2.1.2) sağlamaktadır (Boulesteix, Janitza, Kruppa ve König, 2012).

#### **1.4. META-SEZGİSEL ALGORİTMALAR**

Alanı ve konusu ne olursa olsun incelenen sistemin girdi ve çıktıları arasındaki ilişkilerin ortaya konması, sistemin çalışmasını/performansını etkileyen faktörlerin tespit edilmesi ve geleceğe ilişkin sağlam öngörülere sahip olabilmek için tüm değişkenlerin modellere aktarılabilmesi büyük önem arz etmektedir. Bu kapsamda ve özellikle modern bilim çağında, olguların nesnel olarak incelenmesi ve gözlemlere bağlı olarak sistemlerin modellenmesinde kullanılan gelişmiş matematiksel araçların kullanımı bilim pratiğinin temeli haline gelmiştir. Sistem/süreç optimizasyonu da söz konusu modelleme pratiğinin yoğun olarak kullanıldığı finans ve mühendislik gibi uygulama alanlarında önemli bir yere sahiptir. Optimizasyon, bir sistem içerisinde bulunan ve sistem çıktılarını etkileme kapasitesine sahip değişkenlerin çeşitli alternatifleri arasından en etkin olanı/olanları seçmek olarak tanımlanabilir.

Modern optimizasyon çalışmalarının başlangıcı, matematiksel modellerin yoğun olarak kullanıldığı 1940'lı yıllara dayanmaktadır ve bu çalışmalar halen, başta bilgisayar bilimi olmak üzere gelişen araçlar ve matematiksel yöntemler ile paralel şekilde gelişmeye devam etmektedir. Sistemler ile ilgili toplanan bilgi miktarında ve bu bilgilere bağlı olarak ortaya çıkan çözüm alternatifi sayısındaki muazzam artış, klasik yöntemlerin yetersiz kalmasına yol açmış ve son 30-40 yıllık dönem içerisinde farklı uygulamaların geliştirilmesi ihtiyacını doğurmuştur. Meta-sezgisel algoritmalar gibi yaklaşımlar, gün geçtikçe daha karmaşık hale gelen sistem modellerinin incelenmesinde ve karar verme süreçlerine destek amacıyla kullanılmaya başlanmıştır.

Bu yöntemler arasında meta-sezgisel algoritmalar çeşitli açılardan farklılıklar göstermektedir. Öncelikle her zaman en iyi çözümü garanti etmemektedirler. Doğal süreçlerden esinlenmiş bu yöntemler, bu süreçlerde var olan ve en iyi çözüm yerine en iyiye yakın çözümlerin kabulünü de içeren mekanizmalara sahiptir. Nasıl ki evrimsel

süreçte canlıların kararlarının ötesinde kontrol edilemeyen dışsal faktörlere dayalı etkenler mevcutsa, bu yöntemlerde de en iyi çözüme ulaşma garantisi vermeyen ancak kabul edilebilir çözümler sunan mekanizmalar bulunmaktadır. Bunu, klasik optimizasyon yöntemleri ile meta-sezgisel algoritmalar arasında bir değiş-tokuş olarak görmek mümkündür. Klasik optimizasyon teknikleri, doğru modelleme yapıldığı takdirde en iyi çözümü verme garantisi verirler ancak matematiksel modellemeye uygun olması için ele alınan problemlerde çok sayıda kısıt kullanırlar ya da sistem üzerinde çoğu zaman gerçekçi olmayan varsayımlarla çalışırlar. Olabildiğince gerçeğe yakın olarak modellenmeleri durumunda ise modern bilgisayarların bile yeterince kısa sürede gerçekleştiremeyeceği kadar çok sayıda işleme ihtiyaç duyulabilir ve model hatalarına karşı fazlasıyla hassastırlar. Meta-sezgisel algoritmalar ise, olabildiğince az kısıt ve varsayım ile oluşturulabilirler ve çözüm uzayında alternatif çözümler arasındaki farklılara bağlı olarak tarama yaparak en iyi olmasa bile uygun bir çözümü sunma kapasitesine sahiptirler.

Meta-sezgisel algoritmalar, yalnızca problemin modellenmesinde değil, analizin bir parçası olarak sınıflandırma, değişken ve parametre seçimi gibi işlemler için de kullanılabilir. Bu çalışmada da meta-sezgisel algoritmalar, kullanılan öğrenen yöntemin performansını artırmak için değişken seçimi (bkz. Bölüm 2.2.1) yapılması amacıyla uygulanmıştır.

#### **1.4.1. Biyo/doğa-ilhamlı Hesaplama**

Doğal süreçlerden ilham alan meta-sezgisel algoritmalar arasında en büyük ve önemli grubu, biyolojik organizmaların zekâ, öğrenme ve adapte olma özellikleri üzerine kurulu yöntemler oluşturmaktadır. Bu yaklaşımlar temelde, organizmaların kendileri için en iyi olanı yapma ve bilgiyi aktarma mekanizmalarını taklit edilmesiyle oluşturulurlar. Örneğin, parçacık sürü optimizasyonu (PSO) adı verilen bir meta-sezgisel (Kennedy, 2011), sürü halinde hareket eden karınca, arı gibi hayvanların, yiyecek bulma, kendilerini savunma vb. doğal eylemleri gerçekleştirirken başta sürü olmak üzere kendileri için en iyi seçeneklere yönelimleri ve süreçte edindikleri bilgiyi birbirlerine aktarımları esin kaynağı olmuştur. DVM'nin kredi risk sınıflandırma

performansını artırmaya yönelik olarak parametre seçimi için PSO kullanım örneği Danenas ve Garsva (2015)'in çalışmasında görülebilir.

Biyo-ilhamlı algoritmalar arasında kuşkusuz en popüler olanı, doğal seçim süreci içerisinde üreme, kalıtım, mutasyon gibi mekanizmaların taklit edildiği GA'dır. Biyo-ilhamlı algoritmaların kullanım boyutunu araştırdığı çalışmada Kar (2016), Scopus<sup>6</sup> veri tabanında bulunan makaleler arasında, başlığında GA bulunanların oranının %24.11 olduğunu belirtmiştir. Birinci sırada, %63.04 oranı ile SA bulunmaktadır (SA makine öğrenme kategorisinde yer alan bir yöntemdir). GA ve SA'yı, %8.71'lik oran ile PSO takip etmektedir.

#### **1.4.2. Meta-sezgisel Algoritmaların Kredi Değerlendirme Uygulamaları**

Disiplinler-arası bir alan olarak Yöneylem Araştırması, çeşitli karar verme süreçlerine etkin çözümler getirmeyi amaçlamaktadır ve uzun yıllardır lojistik, ulaşım, imalat, üretim, sağlık bilimleri ve iletişim alanlarında uygulama çalışmaları yürütülmektedir. Soler-Dominguez, Juan ve Kizys (2017), diğer alanlar ile karşılaştırıldığında, veri analiz yöntemlerinde ve finansal sistemler, piyasalar, kurumlar ve varlıklarda yaşanan muazzam gelişim ve değişimlere rağmen uygulamalı finans alanında karşılaşılan problemlerin çözümüne yönelik yeterli ilginin olmadığından söz etmişlerdir.

Ancak son dönemlerde, özellikle sezgisel/meta-sezgisel yaklaşımların kullanımını konusunda artışa geçen bir ilgiden bahsetmek mümkündür. Finansal karar destek araştırmalarında mevcut gelişmeleri aktardıkları çalışmalarında Zopounidis, Doumpos ve Niklis (2018), standart finansal karar modellerinin artan finansal veri ve daha gerçekçi modelleme ihtiyacı karşısında karşılaştığı hesaplama zorluklarının meta-sezgisel algoritmalar ile giderilmesinin güncel bir eğilimi yansıttığını ifade etmişlerdir. Bu eğilimi, kredi değerlendirme çalışmalarında da görmek mümkündür. Bu çalışmalarda meta-sezgisel algoritmalar çok büyük oranda, değişken seçimi yapmak amacıyla kullanılmaktadır.

---

<sup>6</sup> Scopus hakem denetimli bir literatür (bilimsel dergiler, kitaplar ve konferans raporları) özet ve atıf veri tabanıdır.

Zhang, Hifi, Chen ve Ye (2008), GP ve DVM yöntemlerini hibritleyerek, her bir yöntemin ayrı ayrı ve KA ve LR gibi yöntemlerin ulaştığı tahmin başarısından daha iyi sonuçlar elde etmeyi başarmışlardır. Benzer şekilde Zhang, Leung ve Ye (2008), üç aşamalı bir hibrit model kurmuş ve tahmin performansında artış sağlamışlardır: Birinci aşamada GA ile değişken seçimi yapılmış, ikinci aşamada k-ortamalar algoritması ile veri kümelerinde gürültü<sup>7</sup> özelliği taşıyan ve öğrenme performansını azaltan gözlemler analiz dışına çıkartılmış ve son aşamada KA kullanılarak öğrenme sağlanmıştır. Kozeny (2015) de GA kullanılarak başarılı şekilde kredi derecelendirmesi yapılan bir başka çalışma örneğidir.

Chen ve Li (2010) DVM ile sınıflandırma öncesi çözüm uzay boyutunu azaltmak amacıyla veri kümesinde DAA, KA, bir bulanık yaklaşım örneği olan kaba kümeler teorisi gibi yöntemler ile veri ön-işleme gerçekleştirmiş ve tahmin performansında artış sağlamışlardır. Huang, Chen ve Wang (2007), hibritlenmiş DVM ile GA yöntemiyle değişken seçimi ve parametre optimizasyonu gerçekleştirmiştir. Oreski, Oreski ve Oreski (2012), GA kullanarak SA sınıflandırma performansını artırmayı başarmışlardır ve GA ile gerçekleştirilen değişken seçimi sayesinde oluşturulan alt değişken kümesinin bilgi kazanımı (değişkenler arasında en ayırt edici olanı belirlemek için kullanılır ve veri kümesindeki her bir özellik için ölçülür), kazanç oranı (bilgi kazancı değerlerinin, değişkenlerin veriyi ayırma bilgisine oranlanarak normalize edilmesi), Gini endeksi ve korelasyon gibi geleneksel değişken seçimi yöntemleri ile karşılaştırıldığında daha iyi sonuç elde ettiklerini bildirmişlerdir (söz konusu yöntemlere ilişkin ayrıntılı bilgi için bkz. Kaynar, Arslan, Görmez ve Işık (2018)). Bir başka çalışmada da, bilgi kazanımı, kazanç oranı, Gini endeksi ve korelasyon yöntemleri ile hibritleştirilmiş GA ile yapılan değişken seçiminin, SA'nın tahmin performansını artırdığı görülmüştür (Oreski ve Oreski, 2014). Koutanaei, Sajedi ve Khanabaei (2015) çalışmalarında, temel bileşenler analizi (TBA), GA ve bilgi kazanım oranı gibi yöntemler ile değişken seçimi yapılmış, bu yöntemlerin parametre seçimleri DVM ile gerçekleştirilmiş ve en uygun parametreler ile seçilen alt veri kümesi, çeşitli sınıflandırma yöntemleriyle analiz edilmiştir. Yazarlar, önerdikleri

---

<sup>7</sup>Analiz edilen veri kümesinde gözlenen ve rassal yapıya sahip düzensizliklere gürültü adı verilmektedir. Belirli bir örüntüye sahip olmadıklarından modeller tarafından tespit edilmeleri mümkün olamamaktadır.

yöntemin işlevsel ve güçlü bir hibrit model olduğunu belirtmişlerdir. Šušteršič, Mramor ve Zupan (2009), bir Slovenya bankasının 581 müşterisinin borç verileri üzerinde LOGIT ve SA ile yaptıkları sınıflandırma çalışmasında TBA ve GA ile değişken seçimi uygulayarak model performanslarını artırmışlardır.

Türkçe çalışmalar incelendiğinde, kredi değerlendirme üzerine ciddi bir literatür boşluğu bulunduğu anlaşılmaktadır. Mevcut çalışmalarda da yöntem tercihinin analitik hiyerarşi süreci olduğu görülmektedir (Girginer, 2008; İç ve Yurdakul, 2000). Finans alanında yapılan çeşitli çalışmalarda GA kullanım örnekleri bulunsa da bunlar, portföy oluşturma konusuna odaklanmışlardır (Keskintürk, 2007; Özdemir, 2011; Yakut ve Çankal, 2016; Zeren ve Baygın, 2015).

### **1.4.3. Meta-sezgisel Algoritmaların Diğer Finans ve Mühendislik Alan Uygulamaları**

Meta-sezgisel algoritmaları arasında kredi değerlendirme çalışmalarında en sık karşılaşılanı GA'dır. Ancak Soler-Dominguez ve diğerleri (2017)'nin belirttiği gibi kredi risk değerlendirme çalışmalarında meta-sezgisel yaklaşımlar, klasik finans problemlerinin çözüm çalışmalarına nazaran daha az incelenmiş ve yöntem tercihleri birkaç tanesi üzerinde yoğunlaşmıştır. Örneğin kredi değerlendirme çalışmalarında uygulama örneğiyle karşılaşılmayan TB, başta mühendislik olmak üzere başka alanlardaki optimizasyon çalışmalarında kendisine fazlasıyla yer bulmaktadır.

Meta-sezgisel algoritmaların kullanım alan ve konularına yönelik yakın tarihli tarama çalışmalarından alınan bilgiler, Tablo 2'de özetlenmiştir. Algoritmaların kredi değerlendirme yazınında genellikle, istatistiksel ve makine öğrenme yöntemleriyle hibritleştirilerek kullanıldığı görülmektedir. Ancak her yöntemde olduğu gibi bu meta-sezgisel algoritmaların da tek başlarına eksik kaldıkları yönler bulunabilmektedir (tez çalışmasında kullanılan GA ve TB yöntemlerinin güçlü ve zayıf yönleri için bkz. Bölüm 2.2.5). Algoritmaların birlikte ya da iç içe kullanılmasıyla veya algoritma bileşenlerinde analizlere uygun değişiklikler yapılmasıyla zayıf yönler bertaraf edilerek daha başarılı bir optimizasyon sürecine ulaşmak mümkün olabilir. Bilindiği kadarıyla kredi değerlendirme yazınında meta-sezgisel algoritmaların bu tür birlikte

kullanım ya da mekanizma iyileştirme örnekleri bulunmamaktadır. Tez çalışmasında önerilen GA-TB birlikte kullanım örneği, bu açıdan özel bir niteliğe sahiptir.

**Tablo 2:** Meta-sezgisel algoritma kullanım örnekleri

Kaynak	Yöntem	Çalışma Alanı	Çalışma konusu
Chaparala ve Sajja (2018)	Genetik Algoritma, Tavlama Benzetimi	Mühendislik	Çok-amaçlı karar verme
Soler-Dominguez ve diğerleri (2017), Tablo 2	Genetik Algoritma	Finans	Portföy Optimizasyonu, endeks takibi, geliştirilmiş endeks takibi, kredi-risk optimizasyonu, varlık seçimi, proje çizelgeleme
	Tavlama Benzetimi		Portföy Optimizasyonu, endeks takibi, proje çizelgeleme
	Tabu Arama		Portföy Optimizasyonu, proje çizelgeleme
	Diferansiyel Evrim Algoritması		Portföy Optimizasyonu, endeks takibi,
	Genetik Programlama		Varlık seçimi
	Sürü/Koloni Optimizasyonu		Portföy Optimizasyonu, kredi-risk optimizasyonu, proje çizelgeleme
Kar (2016)	Genetik Algoritma	Mühendislik	Gezgin satıcı, sıralama, çizelgeleme, dengeleme problemlerinde, çok-amaçlı/kriterli karar verme ve kısıtlandırılmış optimizasyon problemlerinde maksimizasyon/minimizasyon
Aguilar-Rivera, Valenzuela-Rendón ve Rodríguez-Ortiz (2015)	Genetik Algoritmalar	Finans	Dolandırıcılık tespiti, arbitraj, iflas (temerrüt) tahmini, nakit yönetimi, kredi portföyleri, kredi değerlendirme, temel analiz, tahminleme, endeks takibi, piyasa simülasyonu, portföy optimizasyonu, piyasada işlem yapma

Çalışmanın amacı, parametrik olmayan yöntemler yazınında çeşitli çalışmalarda kullanım örnekleri olan RO yöntemi ile gerçekleştirilecek öğrenme süreci performansının diğer çalışmalar ile kıyaslanabilecek düzeyde geliştirilmesidir. RO yöntemi, tahmin çalışmalarında oldukça başarılı bir yöntem olarak görülmektedir. Başvurulan iyileştirme yönteminin ilk kısmını, kredi değerlendirme yazınında çok sayıda kullanım örnekleri bulunan GA oluşturmaktadır. İkinci kısım ise, mevcut yazında uygulama örneğine rastlanmamış TB algoritmasından oluşmaktadır. İki yöntemin birleştirilmesiyle yeni bir algoritma şeması oluşturulacaktır. İkinci bölümde, söz konusu yöntemlere ilişkin teknik bilgiler ve çalışma prensipleri sunulmuştur.

## **İKİNCİ BÖLÜM**

### **TÜKETİCİ KREDİ BAŞVURULARININ DEĞERLENDİRİLMESİNDE RASTGELE ORMANLAR YÖNTEMİ VE META-SEZGİSEL ALGORİTMALAR İLE İYİLEŞTİRME ÖNERİSİ**

Çalışmada ele alınan veri kümesinin analizi, iki aşamadan oluşmaktadır. Öncelikle, Bölüm 2.1.2’de tanıtılmış olan RO yöntemi kullanılacak, sonuçlar toplandıktan ve incelendikten sonra, Bölüm 2.2.5’te tanıtılmış olan BA-RO melez yöntemi ile elde edilecek iyileştirilmiş performans sonuçları ile karşılaştırılması yapılacaktır. Tüm analizler ve algoritmalar, R Studio (Team, 2015) ara yüzü ile R (sürüm 3.2.3) istatistiksel öğrenme programında (R, 2015) gerçekleştirilmiştir.

Bu bölümde RO, GA, TB yöntem ve algoritmaları ile ilgili temel konulardan bahsedilecektir. GA-TB ile oluşturulmuş BA dizaynı, parametreleri ve algoritmanın yardımıyla gerçekleştirilecek değişken/nitelik seçimi yöntemine ilişkin bilgilerden sonra bölüm, kullanılan ve önerilen yöntemlerin performanslarının değerlendirilmesi için başvuru ölçütlerin tanıtılmasıyla son bulmaktadır.

#### **2.1. AĞAÇ TABANLI YÖNTEMLER**

RO, denetimli makine öğrenme yöntemleri arasında yer alan ve hem sayısal (regresyon) hem de sayısal olmayan (sınıflandırma) bağımlı değişkenli sistemlerde kullanılabilen ağaç tabanlı bir yöntemdir. Ağaç tabanlı yöntemler, temelde, tahmin değişken uzayını ayrıştırarak bölgelere ayırırlar. Ağaç tabanlı yöntemler, veri kümesini “eğer-öyleyse (if-else)” önermeleriyle bölerek, sayısal bağımlı değişken durumunda tahmin hata kareleri toplamını; sınıflandırma durumunda hatalı sınıflandırma oranını en küçüklemeye çalışmaktadır.

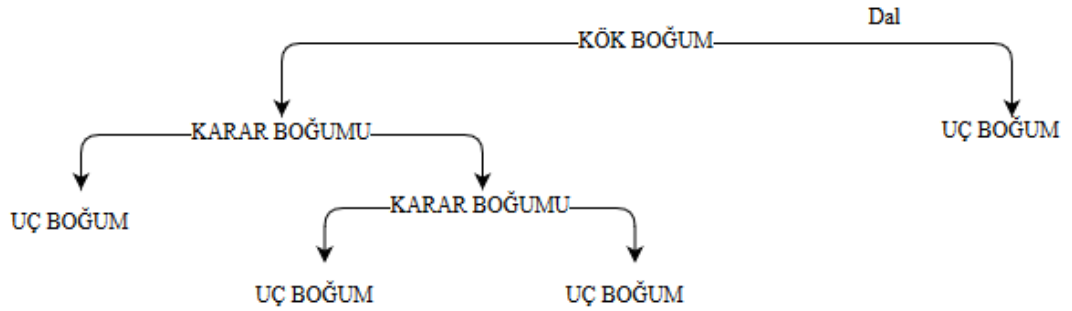
Kredi değerlendirme problemi bir sınıflandırma problemi olduğu için bu bölümde tanıtılacak olan yöntemler, sınıflandırma yöntemleri olarak ele alınacak ve örneklendirilecektir. Ancak performans kriterleri hariç genel yaklaşımların, regresyon ortamında da geçerli olduğunu söylemek mümkündür.

Öncelikle KA’ndan bahsedilecek, ardından da RO yöntemine geçilecektir.

### 2.1.1. Karar Ağaçları

KA'nın genel yapısı Şekil 1'de gösterilmiştir. Kök ve karar boğumları, ağacı iki ya da daha fazla boğuma bölen değişkenleri ifade etmektedir. Bu işleme ayırma (bölme) adı verilir. Ayırma sonrası oluşan dal ile bir sonraki boğum noktasına ulaşılır. Daha fazla ayırma yapılmayan boğum noktalarına uç boğum (yaprak) adı verilir. Yapraklar karar bölgelerini temsil etmektedir. Ayırmalar sonucunda ortaya çıkan karar bölgelerinde, bağımlı değişkenin kategorik değer aldığı durumda hangi sınıfa ait olduğuna, sayısal değer aldığı durumda ise alması gereken sayısal değer ne olduğuna karar verilir.

Şekil 1: Karar ağacı bileşenleri

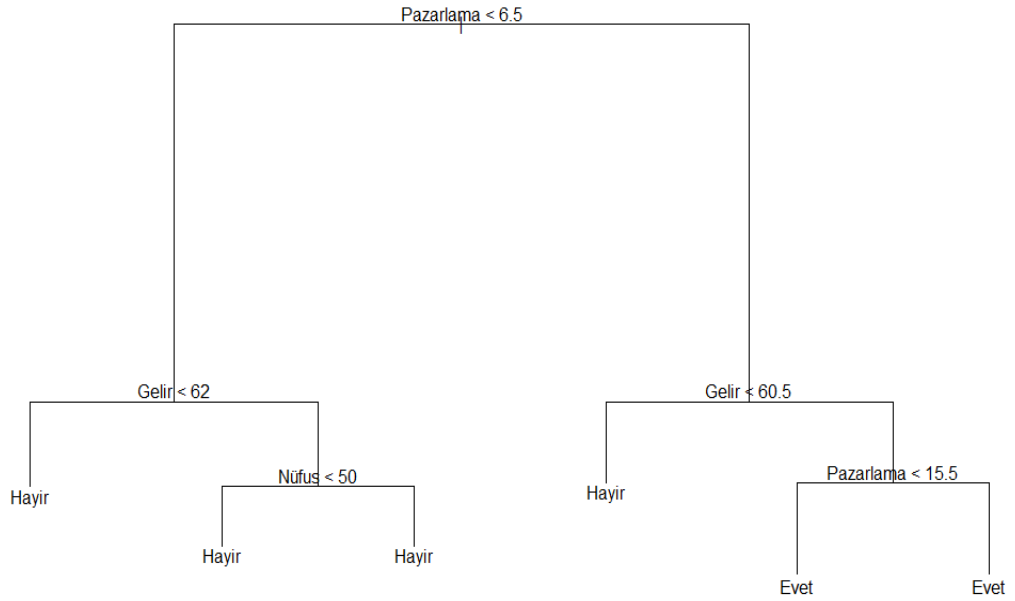


Şekil 2'de, sınıflandırma için kullanılan basit bir KA örneği gösterilmiştir. Şekilde “Evet” ve “Hayır” değerleri, herhangi bir bölgede yapılmış olan araç satışının yüksek olup olmadığına ilişkin sınıflandırma değerleridir. KA, okumalarının ve yorumlanmalarının kolay olması dolayısıyla tercih edilebilmektedir. Örneğe göre, yeni bir gözlemin, Pazarlama değerinin 6.5'ten küçük, Gelir değerinin 62'den büyük ve Nüfus değerinin 50'den büyük olacağı bilinirse/tahmin edilirse, satışların düşük olacağı söylenebilir.

KA'nın sağladığı bir diğer bilgi, hangi tahmin değişkeninin, bağımlı değişken üzerinde daha belirleyici olduğudur. Şekil 2'deki örnekte, Pazarlama değişkeni, ağacı en tepeden kesmektedir. Bu, söz konusu değişkenin, diğer tahmin değişkenlerine göre satışlar üzerinde daha etkili olduğunu işaret etmektedir.



**Şekil 2:** Bir karar ağacı örneği



Veri kaynağı: James, Witten, Hastie, ve Tibshirani, 2017.

$X_1, X_2, \dots, X_p$  1'den  $p$ 'ye tahmin değişkenleri olmak üzere, KA oluşturma adımları şu şekilde özetlenebilir:

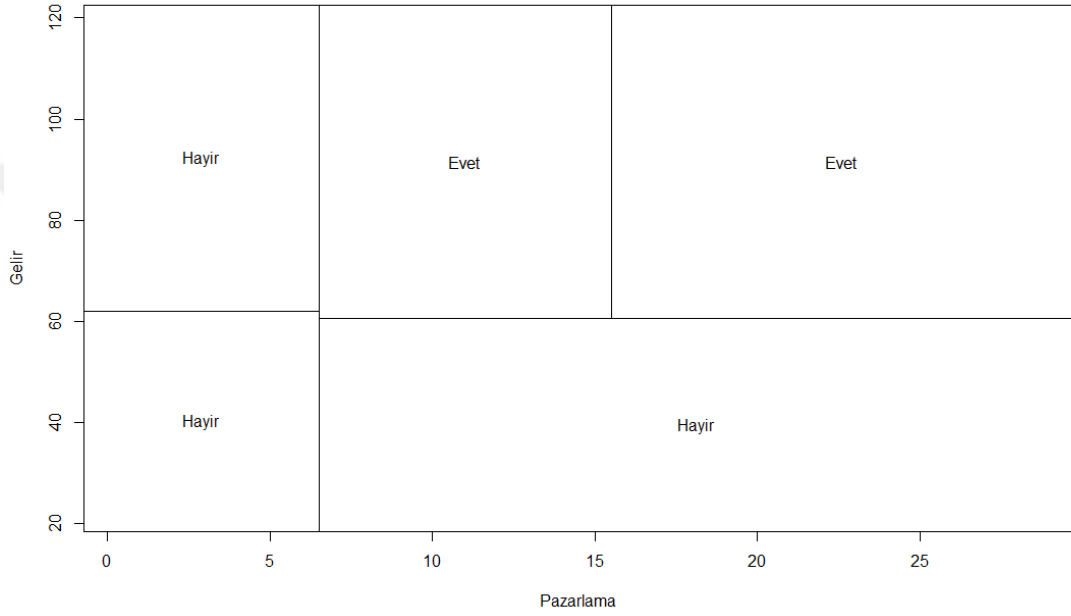
- Adım 1: Her  $X_j, j = 1, \dots, p$  için, bağımlı değişken veri uzayında sınıflandırma hata oranını en küçükleyen bir  $s$  kesim noktası belirle,
- Adım 2: Tüm  $X_{j,s}$  noktaları arasından en küçük hatayı veren ayırmayı seç,
- Adım 3: Durma kriteri sağlanana kadar, ayrılmış veri kümesinin her bir tarafı için, Adım 1 ve Adım 2'yi tekrarla.

Durma kriteri, örneğin oluşturulan her bölgede belli bir sayıdan fazla gözlem kalmaması olabilir. Şekil 3'te, önceden bahsedilen örneğin iki tahmin değişkenli ağaç modelinin ayırma örneği gösterilmiştir.

KA, basitlikleri ve kolay anlaşılabilir olmaları ile tercih ediliyor olsa da yukarıdaki adımlardan da anlaşılacağı üzere sıklıkla, iyi bir sınıflandırma sağlayan “en iyi” karar bölgelerine ulaşamamaktadırlar. Bunun temel sebebi, yöntemin “hırslı” bir

yöntem olması, mevcut ayırma seçenekleri arasındaki en iyiyi seçmekle beraber, ileride olası farklı ve daha başarılı ayırma kombinasyonlarını dikkate almamasından kaynaklanmaktadır. Buna göre, ilk ayırmayı sağlayan tahmin değişkeni diğer tahmin değişkenlerinin etkisine göre ağır basacak ve olası daha büyük etkilerin kaçırılmasına sebep olacaktır.

**Şekil 3:** İki boyutlu veri uzayında KA ayırması ve karar bölgelerinin oluşumu



Veri kaynağı: James ve diğerleri, 2017.

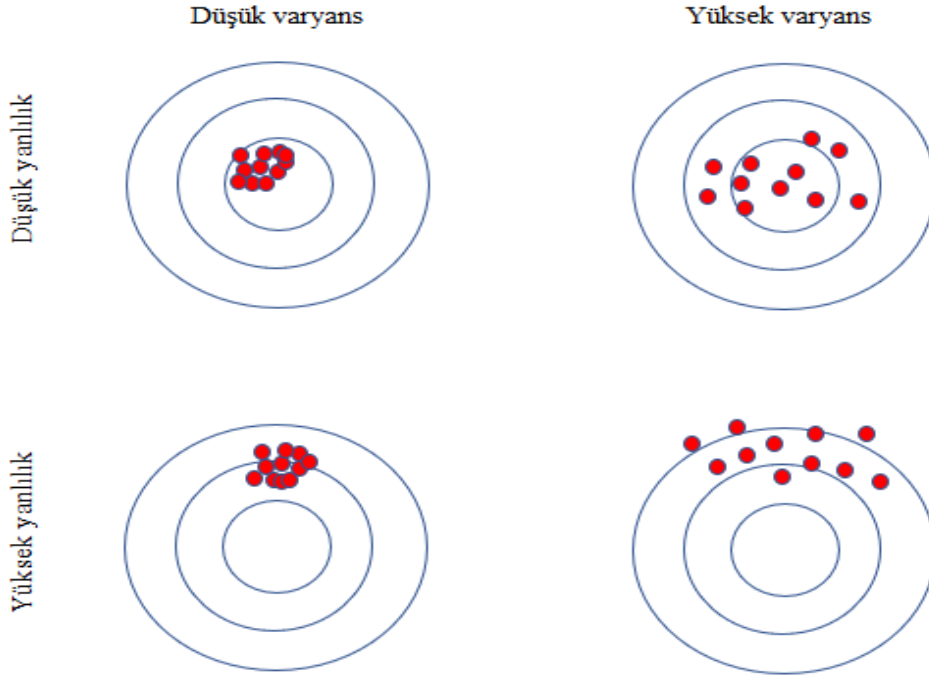
KA yönteminde göz önüne alınması gereken bir diğer husus da öğrenme sürecinin, eğitim için kullanılan veri kümesine aşırı uyum göstermesi durumudur<sup>8</sup>. Karmaşık bir ağaç yerine, daha az sayıda ayırma ile oluşturulan daha küçük bir ağaç, az miktarda artan bir yanlılık karşılığında daha düşük değişkenlik (varyans) elde edilmesini sağlayabilir (James, Witten, Hastie ve Tibshirani, 2013, s. 307).

Yanlılık ve varyans, öğrenen yöntemler ile yapılan tahminleri etkileyen üç temel bileşenden ikisidir. Üçüncü bileşen ise hatadır. Hata, tahmin edilemeyen/öngörülemeyen etkilerden kaynaklandığı ve hiçbir yöntem tarafından modellenemeyeceğinden azaltılamaz bir değerdir. Örneğin hisse senedi getirileri

<sup>8</sup> Aşırı uyum gösterme, eğitim sürecinin eğitim için kullanılan veri kümesindeki değerlere fazlasıyla bağlı kalarak öğrenen yöntemin genelleme yapma yeteneğinin düşmesi durumudur.

zaman serilerinde hata, yeni ve tahmin edilemeyen bilginin (beyaz gürültü) etkisidir. Kredi değerlendirme alanında da hata, modele eklenmemiş/eklenememiş ya da hesaplanmamış/hesaplanması mümkün olmayan bilgi olarak düşünülebilir.

**Şekil 4:** Yanlılık ve varyansın grafiksel gösterimi



Orijinal görsel kaynağı: Fortmann-Roe (2012).

Öte yandan, model ya da model parametreleri seçimi ile varyans ve yanlılık düşürülebilir. Varyans, modelde kullanılan veri kümesi değiştiğinde tahminlerde gözlenen farklılaşma/saçılım ölçüsüdür. Yanlılık ise, karmaşık yapıdaki ilişkilerin gerçekçi olmayan düzeyde basit ya da yanlış olarak modellenmesi sonucu ortaya çıkan tahmin sapmasıdır. Yanlılık ve varyansın tahminler üzerindeki etkisine ilişkin grafiksel bir gösterim Şekil 4'te görülebilir. Düşük yanlılık ve düşük varyansa sahip bir model, farklı veri kümeleri ile eğitilmesi durumunda bile hedef değeri başarılı şekilde tahmin edebilir. Yüksek yanlılık durumunda ise, model varyansı ne kadar düşük olursa olsun hedef değerler doğru olarak tahmin edilememektedir. Kabaca söylemek gerekirse, bir modelin yanlılığının yüksek olması, varyansının yüksek olmasına göre daha istenmeyen bir durumdur. Örneğin düşük yanlılık ve yüksek varyans durumunda, bir hedef değişken üzerinde yapılan farklı tahminlerin bir kısmı

dođru olabilmektedir. Yapılan tahminlerin ortalamasını almak gibi bir yöntem kullanılarak varyans etkisini ortadan kaldırmak mümkün olabilir.

İyi bir model, amaç da göz önünde bulundurulmak kaydıyla hem düşük varyansa hem de düşük yanlılığa sahip olmalıdır. Burada amaçtan kasıt, veri kümesinin analiz edilmesi ile ne yapılmak istendiğidir. Eğer amaç tahmin deđişkenleri ile bağımlı deđişken arasındaki ilişkinin ortaya çıkartılması ve yorumlanması ise, varyansın görece düşük, yanlılığın görece yüksek olduđu yöntemlerin kullanılması problem yaratmayabilir. Ancak amaç bağımlı deđişkenin dođru tahmini ise, varyansın yüksek ve yanlılığın düşük olduđu daha esnek modeller tercih edilmelidir. Esnek modellerde aşırı uyum gösterme olgusunun ortaya çıkıp çıkmadığına da dikkat etmek gerekmektedir. Genel olarak, esnek model tercihi sonucu düşürülen varyans miktarı, artan yanlılık miktarına göre çok daha fazla olabilmektedir.

KA, çeşitli yöntemler ve geliştirmeler ile bahsi geçen problemlerden olabildiğince az etkilenecek şekilde iyileştirilebilir. En basit iyileştirme yöntemi ağaç budamasıdır.

#### **2.1.1.1. Ağaç Budaması**

Ağaç budaması, KA'nın daha az karar bölgesi oluşturarak öğrenilen veri kümesi ile daha iyi genelleme yapmasını sağlamayı amaçlar. Buna göre, belirlenecek bir ayarlama parametresi ile yaprakların daha az sayıda olması sağlanır ve böylece daha basit bir model ile öğrenme gerçekleştirilebilir. Ayarlama parametresi, her yaprak için bir "ceza puanı" niteliği taşımaktadır. Yaprak sayısı artıkça, ağacın performans ölçütünde (sınıflandırma problemi için hatalı sınıflandırma sayısı) negatif etki göstermekte ve karmaşık bir ağaç yerine daha az sayıda yaprak içeren bir ağaç oluşturulmaya çalışılmaktadır.

Bu şekilde oluşturulan bir ağaç ile veri kümesi üzerinde daha iyi bir öğrenim ve tahmin performansında artış sağlanabilmektedir. Ulaşılan ağaç, her ne kadar ilk oluşturulan ağaçtan daha iyi olsa da KA'nın maruz kaldığı yüksek deđişkenlik durumundan etkilenmeye devam etmektedir. Bu problemi gidermek için, torbalama ile ağaç oluşturma yöntemi kullanılabilir.

### 2.1.1.2. Torbalanmış Ağaçlar

Torbalanmış ağaçlar veri kümesi içerisindeki gözlemlerden tekrarlı seçime de izin verilerek, farklı eğitim kümeleri oluşturulması ve çok sayıda KA oluşturulması ile elde edilir. Sınıflandırma tahmini yapılırken de torbalanmış ağaçların içerisinde tüm KA'nın tahminleri arasından, en çok tekrar edeni seçilir. Torbalanmış ağaç, çok sayıda eğitimin sonucunda ortalama alınarak oluşturulduğundan, standart KA'nın maruz kaldığı yüksek değişkenliğin azalmasını sağlayacaktır.

Burada dikkat edilmesi gereken, düşürülen değişkenlik ile sağlanan tahmin performansı artışı karşılığında, veri kümesi tahmin değişkenlerinin bağımlı değişken ile olan ilişkisinin anlaşılmasında düşüşe sebep olduğudur. Ayrıca, birden çok ağaçtan oluşan torbalanmış ağaçlarda, KA'nda olduğu gibi bir görsel ilişki şeması da çıkartılamamaktadır. Bunun yerine, Gini endeksi adı verilen ölçütle oluşturulan ve değişkenlerin önemlerini gösteren bir şema oluşturulabilir. Gini endeksinden, Bölüm 2.1.2'de bahsedilecektir.

Torbalanmış ağaçların bir diğer problemi de oluşturulan her yeni ağacın, önceden bahsedilmiş olan tahmin sürecinin baskın bir değişken tarafından belirlenmesi ve aralarında yüksek korelasyon olan ağaçların ortaya çıkma ihtimalidir. RO metodu, bu eksikliğin giderilmesi için kullanışlı bir araçtır.

### 2.1.2. Rastgele Ormanlar

Çalışmada ele alınan kredi veri kümesinin öğrenilmesi ve tahmin edilmesinde de kullanılacak olan RO, torbalanmış ağaçlar yöntemine benzer şekilde oluşturulmaktadır. Ancak bu yöntemde, baskın değişken varlığı durumunda karşılaşılabilecek yüksek korelasyona sahip ağaçlar sorununu gidermek için ağaç dallarında gerçekleştirilecek her yeni ayırma aşamasında tüm değişkenler değil, rassal olarak seçilen ve önceden belirlenmiş  $m$  sayıda tahmin değişkeni dikkate alınır. Bu şekilde oluşturulan her ağacın tahminleri arasından, torbalanmış ağaçlarda olduğu gibi en çok tekrar edeni seçilir. Bu sayede baskın değişkenin birbirine benzer ağaçlar oluşturması engellenmiş olacaktır (Hastie, Tibshirani ve Friedman, 2009, s. 587).  $m$  değeri, RO için performansı etkileyen bir ayarlama parametresidir.

RO oluřturma adımları, minimum gözlem sayısına ulaşmamıř her bir yaprak üzerinde uygulanmak üzere ařağıdaki řekilde gerekleřtirilmektedir:

Her bir  $T_b$ ,  $b = 1 \dots b_{mak}$ . iin:

Adım 1: Tm yapraklardaki gözlem sayısı, belirlenmiř minimum gözlem sayısına ulařtıysa Adım 5'e git,

Adım 2: Toplam  $p$  deęiřken arasından  $m$  tanesini rassal olarak se,

Adım 3: Seilmiř  $m$  deęiřken arasından, eęitim kümesi üzerindeki en iyi ayırmayı yapan deęiřkeni se,

Adım 4: Yapradı iki alt-boęuma ayır, Adım 1'e dön

Adım 5: Aęacı  $T_b$  olarak kaydet, maksimum aęaç sayısına ulařılmıřsa dur, ulařılmamıřsa Adım 1'e dön.

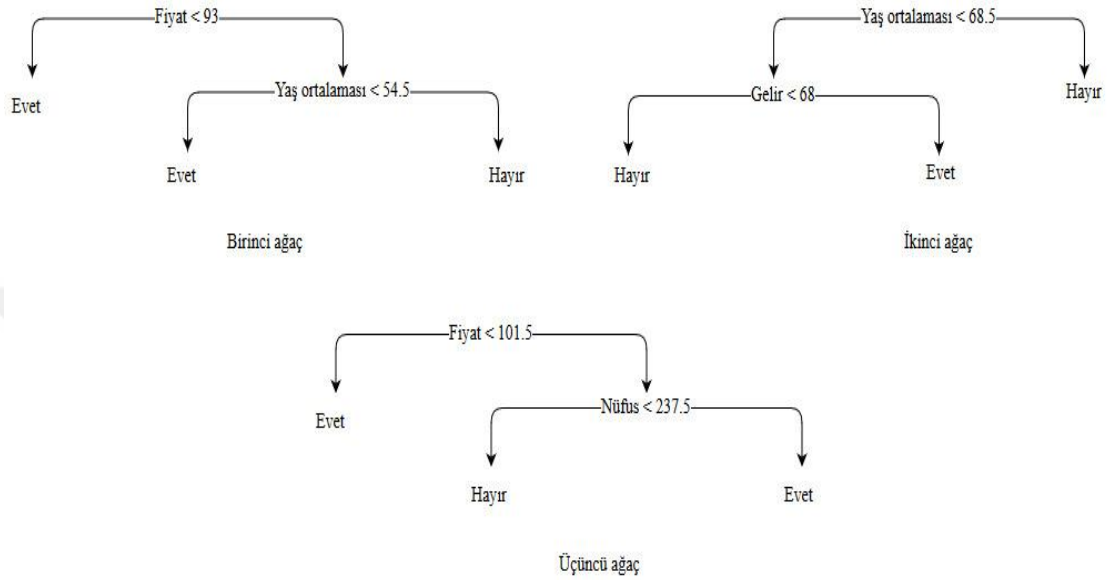
Bu adımlarla oluřturulmuř ormanda nihai tahmin sınıfına, tm aęaçların yaptıęı sınıflandırmalar arasından çoęunluęun belirlenmesi ile ulařılır. Aęaçların oluřturulması iin kullanılan gözlemler, eęitim iin ayrılmıř küme ierisindeki tm gözlemleri arasından ve birbirini tekrar edebilen seimler sonucu oluřturulmaktadır. Bylelikle her aęaç ortaya ıkartılırken, farklı bir eęitim kümesi kullanılmaktadır. Yapraklardaki gözlem sayılarının miktarı, oluřturulan aęaçların derinlięini, bir dięer ifadeyle karmařıklıęını direkt etkileyebilecek bir parametredir.

Oluřturulacak toplam aęaç sayısı, eřitlenmeyi engelleyecek kadar az olmamak řartıyla RO performansını ciddi ölçde etkilememektedir. RO yönteminin ilk geliřtiricisi olan Breiman (2001), RO yönteminin ařırı uyum gösterme olgusuna maruz kalmadıęını belirtmiřtir. ok yüksek aęaç sayılarının model tahmin performansında sorun yaratmadıęını ancak, gereksiz iřlem sürelerine sebep verebileceęini söylemek mümkündür. Gerek yapraklardaki minimum gözlem sayısı gerekse toplam aęaç sayısının ne olması gerektięine iliřkin genel kabul görmüř bir yaklařım bulunmamaktadır.

Hem eęitim kümelerinin hem de ayırma kararı iin kullanılacak olan deęiřken sayısının rassal yapısı, RO yöntemi ile oluřturulan aęaçların birbirinde baęımsız olmasını saęlayan en önemli etkenlerdir. Bylelikle RO yöntemi, yanlılıktan ödünermeden tahmin varyansını düřürmeyi bařarabilmektedir.

Çeşitli mağazalarda satılan çocuk koltuğu miktarlarına ilişkin bir veri kümesinin kullanıldığı ve toplam üç ağaç ile sınırlandırılmış örnek bir RO analizinden elde edilen ağaçlar, Şekil 5’te görselleştirilmiştir.

**Şekil 5:** Rastgele Ormanlar ile çocuk koltuğu satışı analizi örneği



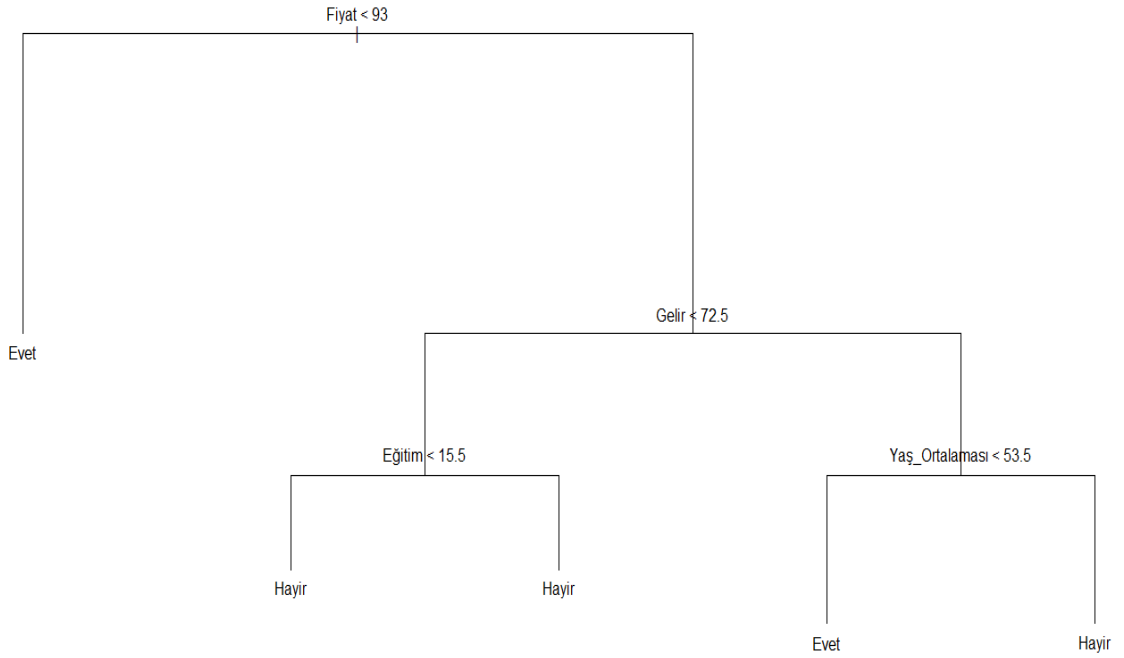
Veri kaynağı: James ve diğerleri, 2017.

Gösterimi basitleştirmek amacıyla, analiz için tüm değişkenler arasından yalnızca “Gelir”, “Nüfus”, “Fiyat”, “Yaş Ortalaması” ve “Eğitim” değişkenleri seçilmiştir. Bu değişkenler, mağazaların bulunduğu bölgeye ait bilgileri içermektedir. Tahmin edilmeye çalışılan değişken *evet ya da hayır* sınıf değerlerine sahiptir ve o mağazada yüksek satış rakamına ulaşıp ulaşılmadığını belirtmektedir. Model yüz gözlem ile eğitilmiştir ve ağaç derinliğini sınırlandırmak amacıyla yapraklardaki minimum gözlem sayısı 30 olarak belirlenmiştir. Ağacın ilk ayrımının gerçekleştiği tepe noktalar, Fiyat değişkeninin baskın bir değişken niteliğinde olduğuna işaret etmektedir. İkinci ağacın ilk ayrımının Yaş Ortalaması değişkeni ile gerçekleştirilmesinde, RO yönteminde her ayrım noktası kararının tüm değişken grubu yerine daha az sayıda ve farklı bir değişken grubunun ele alınması ile verilmesi rol oynamıştır. Bu ağacın ilk ayrım kararında Fiyat değişkeni dışarıda tutularak farklı bir ağacın oluşmasına imkân sağlanmıştır. Benzer şekilde farklılaşma, birinci ve üçüncü ağaçlardaki ikinci ayrımlarda da görülmektedir. Fiyat değişkeninin ayrım noktalarının

iki ağaçta da farklı olmasının sebebi, önceden de belirtildiği gibi her ağacın farklı bir eğitim kümesi ile eğitilmesinden kaynaklanmaktadır.

Şekil 6'da RO yöntemi yerine KA ile analiz yapılması durumunda ortaya çıkacak ağaç gösterilmiştir.

**Şekil 6:** Karar Ağaçları ile çocuk koltuğu satışı analizi örneği



Veri kaynağı: James ve diğerleri, 2017.

KA ile oluşturulan ağaçta da görüldüğü üzere Fiyat değişkeni, beklendiği gibi tüm ağacı en tepeden ayırmaktadır. RO yöntemi ile, baskın olan bu değişkenin etkisi azaltılabilmekte ve hatta yok edilebilmektedir. Şekil 5'teki ağaçlarda tüm değişkenlerin yer almadığı da görülmektedir. Rassal bir süreç olarak ortaya çıkan bu durum, değişkenler arasında bir hiyerarşi oluşturulması için de kullanılabilir. Bu yönüyle RO, Bölüm 2.2.1'de bahsedilecek olan değişken seçimi sürecini de içerisinde (gömülü olarak) barındırmaktadır. Oluşturulan ağaç sayısının az olması, yöntemin olası daha iyi ayırım olasılıklarını araştırmasını engellemektedir. Önceden de



belirtildiği gibi oluşturulacak ağaç sayısının çeşitli değişken kombinasyonlarının oluşmasına yetecek miktarda olması gerekmektedir.

RO ile oluşturulmuş ağaçların, yeni bir gözlem varlığı durumunda yapacağı sınıflandırma tahmini çoğunluk oyuna göre olmaktadır. Örneğin, dört yeni (RO ve KA eğitimine dahil edilmemiş) mağazada çocuk koltuğu satışlarına ilişkin tahmin yapılmak istensin. Mağazaların buldukları bölgelere ilişkin bilgiler Tablo 3'te verilmiştir:

**Tablo 3:** Dört yeni mağaza için çocuk koltuğu satışı tahmin değişkenleri

No.	Gelir	Reklam	Nüfus	Fiyat	Yaş Ortalaması	Eğitim
1	118	12	272	151	43	14
2	65	5	298	125	62	12
3	29	11	335	127	58	12
4	87	9	17	95	65	13

Bu veriler ile KA ve RO ağaçlarının yapacakları tahminler, Tablo 4'te gösterilmiştir:

**Tablo 4:** Dört yeni mağaza için çocuk koltuğu satışı tahminleri

Ağaç/Mağaza No.	1	2	3	4
Birinci ağaç (a)	EVET	HAYIR	HAYIR	HAYIR
İkinci ağaç (b)	EVET	HAYIR	HAYIR	EVET
Üçüncü ağaç (c)	EVET	EVET	EVET	EVET
Çoğunluk Kararı	EVET	HAYIR	HAYIR	EVET
Karar ağacı	EVET	HAYIR	HAYIR	HAYIR
<b>Gerçekleşen Satış</b>	<b>HAYIR</b>	<b>HAYIR</b>	<b>HAYIR</b>	<b>EVET</b>

Tabloda görüldüğü gibi, RO yöntemi %75, KA yöntemi ise %50 tahmin başarısına ulaşmıştır. RO yönteminin her ağacının doğru tahmin vermediğine dikkat edilmelidir. Örneğin birinci ağaç %55 oranında doğru tahminde bulunurken, ikinci ağaç %75, üçüncü ağaç %25 doğru tahmin oranlarına sahiptir. Çoğunluk kararı uygulaması ile toplamda, genel tahmin başarı ortalamasının çok daha üzerine ulaşmak mümkün olmaktadır. Her ağaç diğer ağaçlardan bağımsız olarak oluşturulduğu için RO, gürültü faktörü yüksek bağımsız değişken varlığı durumuna karşı oldukça dayanıklıdır.

Bir ağaç oluşturulurken verilecek her ayırma kararında, salt sınıflandırma hatalarının en küçüklenmesi hedeflenmemektedir. Hem ağacın genelleme yeteneğini

koruması hem de karmaşık olmayan bir yapıya sahip olması için daha basit (safılık derecesi yüksek) bir ağaç elde edilmeye çalışılır. Bu amaçla da Gini Endeksi adı verilen bir ölçüt kullanılmaktadır.

Bir ağacın yapraklarında bulunan gözlemlerin sınıf oranları, yaprakların doğru sınıflandırma başarılarına ilişkin bilgi vermektedir. Buna göre her  $k$  sınıfı için:

$$(1) \quad E = 1 - maks(pro_{nk}),$$

formülü ile hata oranları hesaplanabilir.  $pro_{nk}$ ,  $n$  bölgesinde (yaprağında)  $k$  sınıfına ait gözlem oranını gösterir ve aşağıdaki gibi hesaplanır:

$$(2) \quad pro_{nk} = \frac{1}{N_n} \sum_{x_i \in R_n} I(y_i = k),$$

burada  $N_n$ ,  $n$  bölgesinde ( $R_n$ ) bulunan toplam gözlem sayısı,  $x_i$ ,  $n$  bölgesinde bulunan  $i$ . gözlemi ifade etmektedir.  $y_i$ ,  $i$ . gözlemin bağımlı değişken sınıf değeridir. (1) numaralı formül ile hesaplanan hata oranı yerine kullanılabilir olan Gini Endeksi:

$$(3) \quad G = \sum_{k=1}^K pro_{nk}(1 - pro_{nk}),$$

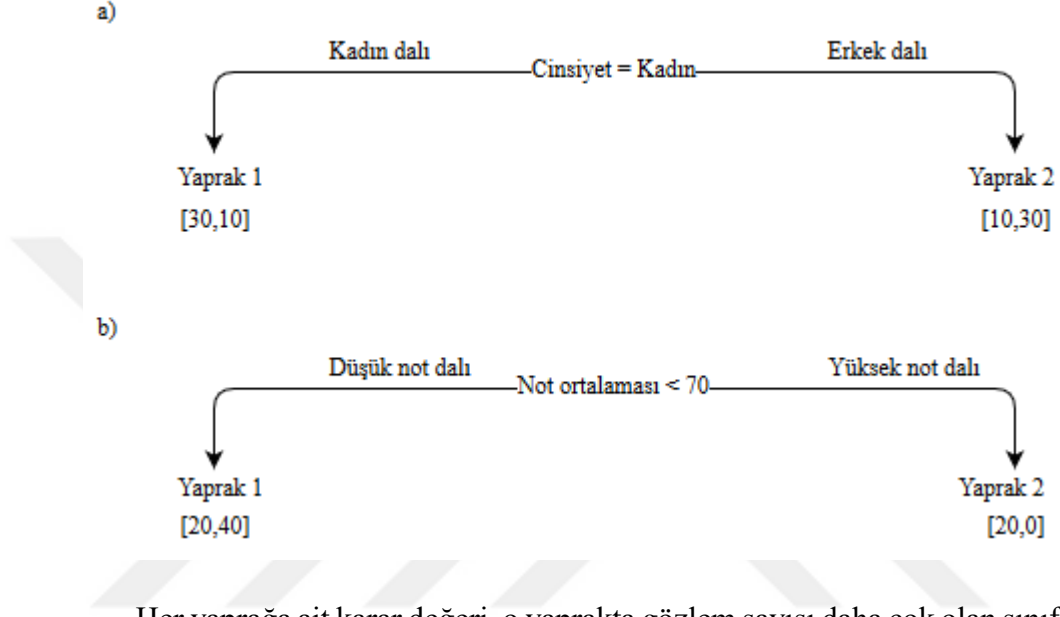
formülü ile hesaplanır. Burada  $K$ , tahmin edilmeye çalışılan toplam sınıf sayısıdır. Gini endeksi, her ayırım anında, olası bölgelerdeki sınıfların ne derece baskın olduğuna ilişkin bir ölçüttür ve ayırım noktaları ile oluşturulacak bölgelerin saflığının hesaplanmasında kullanılmaktadır. Düşük Gini Endeks değeri, yeni oluşturulacak bölgenin tek bir sınıftan yoğun şekilde gözlem içerdiğini belirtir (James ve diğerleri, 2013, s. 312).

### 2.1.2.1. Gini Endeksi Örnek Bir Ayırım Kararı

Toplam 80 öğrencinin olduğu bir sınıfta felsefe dersini alan öğrencilerin tahmin edilmesine yönelik bir ağaç oluşturulmak istensin. Tahmin için iki adet değişken kullanılabilir: Cinsiyet (*kadın*, *erkek*) ve not ortalaması (50-100 arası

sayısal değerler). Sınıfta felsefe dersini alan ve almayan öğrenci sayısı eşittir (birinci sayı alanlar ikinciler almayanlar olmak üzere [40, 40] olarak gösterilsin). Bu senaryoda, ilk ayırım kararı verileceği zaman iki karar değişkeni seçeneği bulunmaktadır. Her iki seçeneğe ilişkin ağaç görselleri Şekil 7’de gösterilmiştir.

**Şekil 7:** Örnek ayırma kararı süreci



Her yaprağa ait karar değeri, o yaprakta gözlem sayısı daha çok olan sınıfa göre belirlenmektedir. Örneğin Şekil 7-a ağacındaki Yaprak 1 için karar değeri felsefe dersini almak ( $30 > 10$ ) olacaktır. Buna göre yaprakların hatalı sınıflandırma oranları  $E$ , formül (1)’den:

- Şekil 7-a Yaprak 1 için:  $1 - \max(30/40, 10/40) = 1 - 3/4 = \%25$ ,
- Şekil 7-a Yaprak 2 için:  $1 - \max(10/40, 30/40) = 1 - 3/4 = \%25$ ,
- Şekil 7-b Yaprak 1 için:  $1 - \max(20/60, 40/60) = 1 - 2/3 = \%33.\bar{3}$ ,
- Şekil 7-b Yaprak 2 için:  $1 - \max(20/20, 0/20) = 1 - 1 = \%0$ ,

olarak hesaplanabilir. Ayırım noktaları ile ulaşılan toplam hata oranları ise her ayırım ile oluşan yaprakların hata oranlarının ağırlıklandırılmış olarak toplanması ile elde edilir; cinsiyet için:  $\%25 * 40/80 + \%25 * 40/80 = \%25$ ; not ortalaması için  $\%33.\bar{3} * 60/80 + \%0 * 20/80 \approx \%25$ . Görüldüğü gibi hata oranları eşittir. Buna göre her iki ayırım arasında fark eden bir şey olmayacaktır.

Yaprakların saflık düzeylerine bakıldığında ise daha farklı bir sonuçla karşılaşılacaktır:

- Şekil 7-a Yaprak 1 için:  $30/40 * (1-30/40) + 10/40 * (1-10/40) = \%37.5$ ,
- Şekil 7-a Yaprak 2 için:  $10/40 * (1-10/40) + 30/40 * (1-30/40) = \%37.5$ ,
- Şekil 7-b Yaprak 1 için:  $20/60 * (1-20/60) + 40/60 * (1-40/60) = \%44.4$ ,
- Şekil 7-b Yaprak 2 için:  $20/20 * (1-20/20) + 0/20 * (1-0/20) = \%0$ ,

Buna göre en yüksek saflığa sahip yaprak, not ortalamasının 2. yaprağı olmuştur (en düşük endeks değeri ile). Ayrım noktalarının ağırlıklandırılmış Gini Endeksi değerleri; cinsiyet için:  $\%37.5 * 40/80 + \%37.5 * 40/80 = \%37.5$ ; not ortalaması için  $\%44.4 * 60/80 + \%0 * 20/80 = \%33.3$  olacaktır. Böylelikle bu noktada ayrımın not ortalaması değişkeninin 70 değeri ile yapılması tercih edilecektir<sup>9</sup>.

Yaprak saflığının ölçüt olarak alınmasının, ağaç oluşumunda ilginç bir etkisi bulunmaktadır. Örneğin Şekil 6'daki ağaçlarda, "Eğitim <15.5" kararı ile oluşan yaprakların her ikisinde de aynı sonuca ulaşılmaktadır. Bunun sebebi söz konusu ayrımın, hatalı sınıflandırma kriteri yerine saflık oranına bağlı olarak oluşturulması ve böylece yüksek saflığın olduğu bölgelere düşen yeni gözlem tahmin doğruluklarından daha fazla emin olunmasıdır. Bu etkinin, öğrenme sürecinde genelleme yeteneğinde artış sağlayacağı aşikârdır.

Şimdiye kadar bahsi geçen ve RO yöntemi içerisinde yer alan tüm hesaplamaların yapılması ve kararların verilmesinde, R programının *randomForest* (Liaw ve Wiener, 2002) kütüphanesinde bulunan *randomForest* fonksiyonu kullanılmıştır. Fonksiyon parametrelerinden olan *m* (her yeni ayırma aşamasında dikkate alınacak değişken sayısı) değeri, toplam tahmin değişkeni sayısının karekökü alınarak belirlenmiştir. Minimum gözlem sayısı 1, oluşturulacak ağaç sayısı da 500'dür. RO ile, en iyi yaprak saflığına sahip bir model oluşturularak etkin bir kredi sınıflandırma tahmini yapılması amaçlanmıştır.

---

<sup>9</sup> En iyi kararın verilebilmesi için sayısal değışkende ortaya çıkabilecek tüm ayırım noktalarının denenmesi gerekmektedir. Bu noktalar genellikle birbirini takip eden iki sayısal değerin orta noktası olarak belirlenir.

## 2.2. GENETİK ALGORİTMA- TAVLAMA BENZETİMİ

Pratikte karşılaşılan birçok problem, çok kısıtlı ve basitleştirilmiş modeller ile çözülemeyecek kadar karmaşıktır. Yöneylem araştırmacıları uzun yıllar boyunca, hem ele alınan problemin hesaplanabilir olması gerekliliğinden çok sayıda kısıt ile çalışmak zorunluluğu; hem de fazlasıyla uzun hesaplama süreleri ile karşı karşıya kalmıştır. Zor kombinatoriyal optimizasyon problemleri, çözüm uzaylarının büyüklüğü sebebiyle en iyi çözümün bulunmasının neredeyse imkânsız olduğu ya da sonuca anlamlı bir süre içerisinde ulaşılamayacak problemlerdir. Problem için kesin ve optimum çözüm veren yaklaşımlar, problem karmaşıklıkla işe yaramaz hale gelmektedir (daha fazla bilgi Bölüm 572.2.2’de verilmiştir).

Sezgisel yaklaşımlar, anlamlı bir süre içerisinde en iyi olmasa da iyi bir çözüm bulmak amacıyla dizayn edilmiş yöntemlerdir. Ancak sezgisel yaklaşımların iki temel sakıncası vardır: Öncelikle probleme özel olarak geliştirilirler, çözümlerin başka problemlere uyarlanmaları zor ve hatta imkansızdır. İkinci olarak, tek bir çözümü en etkili şekilde oluşturmak üzere dizayn edilmişlerdir. Bu durum, çok sayıda uygun çözümü olan problemlerde ulaşılabilecek daha iyi çözümlerin araştırılmasına engel olmaktadır (Zäpfel ve Braune, 2010, s. v).

Meta-sezgisel algoritmalar, bahsi geçen sakıncalara karşı başarılı çözümler sunmaktadır. Temel olarak meta-sezgisel algoritmalar, Bölüm 1.4’te de belirtildiği gibi çoğunluğu evrimsel/doğal süreçlerden esinlenerek oluşturulan ve çözüm uzayını etkin şekilde tarayarak, en iyi olmasa da en iyiye yakın bir çözüm sunan ve bunu anlamlı bir süre içerisinde gerçekleştiren yaklaşımlardır. Meta-sezgisel algoritmaların çalışma mekanizmalarının genel bir yapıda olması ve ele alınan probleme kolaylıkla adapte edilebilmesi, son yıllarda çok farklı alanlarda kullanılmalarını sağlamaktadır.

Meta-sezgisel algoritmalar kabaca, *tek-çözüme dayalı* ve *popülasyona-dayalı* yöntemler olmak üzere iki sınıfa ayrılabilir. Tek-çözüme dayalı yöntemlerde (TB, Tabu arama vb.) tek bir çözüm üzerinden çözüm uzayında yinelemeli olarak tarama yapılır. Popülasyona-dayalı yöntemlerde ise çözüm uzayı taraması bir çözüm kümesi üzerinden gerçekleştirilir (GA, PSA, Diferansiyel Evrim Algoritması vb.). Çalışmada yapılan kredi değerlendirme analizi için değişken seçimi gerçekleştirmek üzere GA ve TB sezgiselleri kullanılmıştır. Bu yöntemlere ilişkin ayrıntılı bilgiler, değişken seçimi

stratejilerinden ve problem zorluk sınıflarından bahsedildikten sonra izleyen bölümlerde sunulmuştur.

### 2.2.1. Değişken Seçimi

Değişken seçimi, bir değişken kümesi içerisinde, oluşturulan modellere dahil edilecek ve bağımlı değişken ile ilişkili olduğu düşünülen daha az sayıda bağımsız değişken içeren bir alt küme oluşturulma sürecidir. Değişken seçimi esas olarak, bilgi sağlamayan ya da alakasız tahmin değişkenlerinin modellerden çıkartılmasına odaklanmaktadır (Kuhn ve Johnson, 2013, s. 488). Çok sayıda değişken içeren ve genellikle, bilimsel teorilerden bağımsız olarak yeni modeller kurulması ya da test edilmesi aşamalarında oluşabilecek karmaşıklık problemlerinin önüne geçmek için veri madenciliği/analizi çalışmalarında sıklıkla başvurulan bir süreçtir. Modellerin basitleştirilmesi ile sonuçlardan yola çıkılarak yapılacak istatistiksel çıkarımların gücünün artması, aynı bilgiyi sağlayan değişkenlerin yaratacakları çoklu bağlanım sorununun ortadan kalkması ve işlem süre/maliyetlerinin azaltılması sağlanabilir. Değişken seçiminin analiz süreçlerine yapıları katkıları kısaca beş maddede özetlemek mümkündür (Salappa, Doumpos ve Zopounidis, 2007): (1) Veri kümesindeki gürültü faktörlerini azaltmak; (2) Uygun modellere ulaşmak için gereken işlem maliyetini azaltmak; (3) Sınıflandırma algoritmasının nihai modelinin anlaşılabilirliğini artırmak (bu fayda özellikle yorumlama çalışmalarında büyük katkı sağlamaktadır); (4) Uygulamayı basitleştirmek; ve (5) Modelin güncellenmesine destek olmak. Burada bahsi geçen birçok etken, değişken seçim süreçlerinin veri madenciliği ve yorumlaması çalışmalarında sıklıkla tercih edilmesini sağlamaktadır.

Değişken seçimi yöntemlerini kabaca filtre, sarmal, hibrit ve gömülü teknikler olarak gruplandırmak mümkündür. Filtre tekniği, bu gruplar arasında uygulaması en basit olanı ve en hızlı çalışandır. Bu yöntem ile veri kümesindeki değişkenler, küme analize sokulmadan önce çeşitli seçim kriterlerine bağlı olarak azaltılır. Buna örnek olarak, değişkenler arası korelasyon değerlerine göre yüksek ilişki düzeyine sahip değişkenlerden bir tanesinin temsilci olarak seçilmesi ve diğerlerinin analiz dışı tutulması gösterilebilir. Filtre teknikleri hesaplama maliyetleri açısından çok etkin olsa

da deęişken seçim kriterleri model performans etkinlikleriyle direkt olarak ilişkili deęildir.

Sarmal teknięinde ise bařlangıçta seçilmiş bir deęişken grubu, sınıflandırma yöntemi ile analiz edildikten ve analiz başarısı deęerlendirildikten sonra gruba ekleme ve çıkartılma yapılarak tekrar deęerlendirmeye alınır ve böylece en iyi sonucu veren deęişken grubu belirlenmeye çalışılır. Sarmal teknik filtre teknięine göre çok daha başarılı olsa da sınıflandırma yöntemine ve veri kümesi büyüklüęüne baęlı olarak artan işlem maliyeti ve seçim performansının büyük veri kümelerinde çok başarılı olamaması gibi dezavantajlara sahiptir. Sarmal tekniklerde deęişkenlerin ne şekilde modele ekleneceęi ve deęerlendirileceęine ilişkin temelde üç çeşit yaklaşım bulunmaktadır:

*İleriye Doğru Seçim:* İleriye doğru seçim yönteminde aday deęişkenler modele teker teker eklenir ve istatistiksel olarak anlamlı katkı sağlayanlar seçilir. Anlamlılık, örneęin doğrusal regresyon modelinde deęişken katsayılarının p-deęerleri ya da model tahmin performansında artış olarak belirlenebilir. Seçim prosedürü, model dışarısında kalan hiçbir deęişkenin anlamlı katkısının olmadığı adıma kadar çalışmaya devam eder. İleriye doğru seçim yöntemi, öğrenme literatüründe *açgözlü* olarak adlandırılan yöntemler arasında bulunmaktadır. Açgözlü yöntemler yalnızca mevcut koşulları deęerlendiren, geçmiş koşulları tekrar deęerlendirmeyen ya da gelecekte ulaşılabilir daha iyi sonuçlar için deneme dizaynına sahip olmayan yaklaşımlardır. Hipotez testleri ile yapılan deęişken seçimlerine ilişkin Harrell (2001), bu şekilde otomatikleştirilmiş yöntemlerin tüm istatistiksel tahmin ve hipotez test prensiplerini ihlal ettięini belirtmiştir. Bu sebeple, tahmin hataları, bilgi kriterler vb. seçim kriterleri kullanılması daha uygun olacaktır.

*Geriye Doğru Seçim:* Geriye doğru seçim, ileriye doğru seçim yöntemine benzer şekilde çalışmaktadır. Bu yöntemde model, tüm aday deęişkenlerin eklenmesiyle çalıştırılır ve istatistiksel olarak anlamlı katkı sağlamayanlar tek tek elenir. Geriye doğru seçim, ileriye doğru seçimin maruz kaldığı dezavantajları aynen taşımaktadır.

*Adım Adım Seçim:* Adım adım seçimde, modele yeni deęişken eklenmesi/çıkartılması durumunda tüm deęişken etkileri tekrar deęerlendirilmektedir. Böylelikle daha önceden modele eklenmiş bir deęişken, eęer başka bir deęişken

varlığında anlamlı bir katkı sağlamıyorsa ya da performansı düşürüyorsa modelden çıkartılabilmektedir. Adım adım seçim yöntemi, ileriye ve geriye doğru seçim yöntemlerine göre çok daha uzun işlem sürelerine mal olabilmektedir çünkü en olumsuz senaryoda olası tüm değişken kombinasyonlarını değerlendirmek gerekebilir.

Tez çalışmasında kullanılan meta-sezgisel yöntemler de bu tekniğe bağlı olarak seçim yapılmasını sağlamaktadır. GA ve TB gibi yöntemler ile gerçekleştirilen adım adım seçim ile söz konusu araştırma süresi dezavantajı ortadan kaldırılabilir. Meta-sezgisel algoritmalar ile seçim sürecine ilişkin daha fazla bilgi ilerleyen bölümlerde sunulmuştur.

Filtre ve sarmal tekniklerin avantajlarını kullanmak üzere geliştirilen hibrit modeller, filtre tekniği uygulanmış bir veri kümesine sarmal teknik uygulayarak tahmin başarısı ile işlem hızı arasında denge kurmaya çalışmaktadır (Jin, Jin ve Qin, 2012). Hibrit tekniğine örnek olarak, kredi derecelendirmesi alanında GA ve bilgi kazanımı yöntemlerinin birlikte kullanıldığı Jadhav, He ve Jenkins (2018)'in çalışması gösterilebilir. Son grup olan gömülü teknikler, sınıflandırma algoritmasının içinde bulunan öznitelik seçme algoritmalarıdır (Gümüşçü, Aydilek ve Taşaltın, 2016). Örneğin KA ve RO, eğitimleri sırasında analiz değişkenleri arasında bir hiyerarşi kurarak en iyi sınıflandırma alternatifini belirlemeye çalışırlar (bkz. Bölüm. 2.1).

### 2.2.2. P, NP ve NP-zor Problemler

En iyi öğrenmeyi sağlayan değişken grubunun seçilmesi, polinomsal zamanda çözülemeyen problemler sınıfına girmektedir. Karmaşıklık Kuramı'na<sup>10</sup> göre bazı tip problemlerin çözümü için en etkili algoritmaların çalışma süresi, girilen verinin büyüklüğüne bir polinom cinsinden bağlı olduğu bilinmektedir (buna polinomsal zamanda çalışan algoritma adı verilir). Bu tür problemler P (polynomial; polinomsal) kategorisine girerler ve veri kümesi büyüklüğünden bağımsız olarak etkin bir sürede

---

<sup>10</sup> Karmaşıklık Kuramı'nın temelinde problemleri çözüm zorluklarına göre sınıflandırmak yer alır. Problemleri çözümler ve çözümlenmez olarak sınıflandırmaya çalışan Hesaplanabilirlik Kuramı ve hesaplama modellerinin özellikleri ve sınırlarını araştıran Özdevinim Kuramı ile birlikte Hesaplama Teorisi'ni oluştururlar. Hesaplama teorisi, bilgisayarların işlem sınırlarını matematiksel modeller ile açıklamaya çalışan bir bilgisayar bilimi dalıdır (Maheshwari ve Smid, 2017, ss. 1-3).



çözülebilirler. P problemlere örnek olarak n adet isim içeren bir listenin alfabetik sıraya göre sıralanması veya liste içerisinde belirli bir ismin bulunması verilebilir.

NP (non-deterministic polynomial; belirleyici olmayan polinomsal) problemlerin çözüm süreleri ise problem büyüklüğüne bağlı olarak hızlı bir şekilde artabilmekte ve hatta anlamlı bir süre içerisinde çözüme ulaşmak mümkün olamamaktadır. Bu sınıfa ait problemlerin çözümü zor olmasına rağmen, verilen bir çözümün doğruluğunun kontrolü P sınıfına girmektedir. Örnek olarak verilebilecek Sudoku problemlerinin çözümüne ulaşmak boyut sayısı arttıkça zorlaşmakta ve hatta insan ömrünü aşan süreler mal olabilmektedir. Ancak verilen bir çözümün doğruluğunu kontrol etmek çok kısa süreler içerisinde gerçekleştirilebilir.

NP-zor sınıfına giren problemlerde, verilen çözümün doğru ya da en iyi olup olmadığını anlamakta da problem boyutuna bağlı olarak değişmektedir. Bu sınıfın en bilindik örneği gezgin satıcı problemi<sup>11</sup>. Probleme sunulan çözümün (rotanın) en iyi alternatif (en kısa rota) olduğunun kontrolü için de en az çözüme ulaşmak kadar uzun süreler ihtiyacı duyulacaktır. İşlem süresini kısaltmak için bilgisayar işlemci gücünü artırmak mümkün olabilse de bu yaklaşım, problemin genel hesaplanabilirliği üzerinde sürdürülebilir bir etkiye sahip olamayacaktır. Problem karmaşıklığının giderilmesine yönelik gerçek bir geliştirmeye yalnızca daha iyi bir çözüm algoritması/yöntemi ile ulaşılabilir (Maringer, 2005, s. 41).

Bu çalışmada hedeflenen en iyi tahmin performansını sağlayan değişken grubunun seçilmesi problemi de NP-zor sınıfa girmektedir. 20 değişkenli ve ikili gösterim kullanılan bu çalışmada, olası tüm çözümlerin kombinasyonu,  $2^{20} = 1.048.576$  olacaktır. Her olası en iyi çözümün eğitimde kullanılması için  $t$  kadar süre harcadığı varsayılırsa tüm çözümler arasından en iyisinin araştırılabilmesi için gereken süre  $1.048.576t$  olur. Daha fazla değişkene sahip bir analiz için de söz konusu sürenin üstel olarak artacağı aşikardır. Aynı şekilde bir çözümün en iyi olup olmadığının kararı için de bir o kadar işlem süresi gerekecektir. Süre kısıtlarının olduğu ya da yeterli işlemci gücün bulunmadığı durumlarda, GA ve TB gibi meta-sezgisel algoritmalar yardımıyla olası tüm çözümlerin araştırılmasına gerek kalmadan iyi bir tahmin performansına ulaşmak mümkün olabilir. İlerleyen bölümlerde,

---

<sup>11</sup> Probleme amaç, bir satıcının bulunduğu konumdan başlayarak her hedef noktaya sadece bir kez uğradıktan sonra başladığı konuma dönmesini sağlayacak en kısa turu bulmaktır.

değişken seçimi için kullanılmış olan meta-sezgisel yöntemlerin işleyiş mekanizmaları ve bileşenleri tanıtılacaktır.

### 2.2.3. Genetik Algoritma

GA doğal süreçlerden esinlenerek geliştirilen meta-sezgisel yaklaşımlar arasında en popüler olanıdır. Terim ilk kez, John Holland tarafından yayınlanan bir kitap ile kullanılmaya başlanmıştır (Holland, 1992). GA, evrim sürecinden örnek alınarak, en iyi çözüm arayışını seçim, çapraz birleşim ve mutasyon bileşenleri ile gerçekleştirir. GA'yı daha iyi anlamak için temel bazı kavramların ve operatör adı verilen bileşenlerinin bilinmesi gerekmektedir.

#### 2.2.3.1. Kromozom/birey

GA'da her çözüm, probleme uygun olarak tasarlanmış olan bir *kromozom/birey* (bundan sonra birey kullanılacaktır) tarafından temsil edilmektedir. Her birey, *gen* adı verilen çözüm değerlerinden oluşmaktadır. Oluşturulan genler, bilgisayar programında tek boyutlu bir dizinin hücreleri ile ilişkilendirilmektedir.

Bireylerin alabileceği çözüm değerlerinin ne şekilde kodlanacağı ele alınan probleme göre farklılık göstermektedir. Bu çalışmada, ikili kodlama kullanılmıştır. Buna göre bir değişkenin analize dahil edilip edilmeyeceğine ilişkin karar sayısal olarak iki değer ile tanımlanmıştır; "1", dahil edildiğini, "0" dışarıda bırakıldığını gösterir. Şekil 8'de ikili değerler ile kodlanmış bir birey örneği gösterilmiştir.

**Şekil 8:** İkili olarak kodlanmış bir birey örneği

Değişken No:	1	2	3	4	5	6	7	8	9	10
Birey	0	1	1	1	0	1	0	1	0	1

Seçilmiş değişkenler: 2, 3, 4, 6, 8, 10

Birey kodlamasına göre sınıflandırma tahmininde kullanılmak üzere 10 değişken arasından 6'sı seçilmiştir. Bu seçim, öğrenen yöntemler terminolojisinde değişken/nitelik seçimi olarak geçmektedir. Her çözüm değerinin atandığı numaralar

(değişken numarası/endeksi), farklı bir değişkeni temsil etmekte ve analiz boyunca sabit kalmaktadırlar.

### 2.2.3.2. Popülasyon

GA, çözüm uzayında arama sürecine, önceden belirlenmiş sayıda birey içeren bir *popülasyon* ile başlar. GA'nın performansları açısından, popülasyon büyüklüğü belirleme önemli yer tutmaktadır. Diğer tüm operatörlerin aynı olduğu durumda çok sayıda birey içeren bir popülasyon gereğinden fazla işlem sürelerine sebep olabilir ya da tersi olarak, az sayıda birey, *prematüre yakınsama* adı verilen ve çözüm uzayında yeterli tarama yapamadan başlangıç döngülerindeki çözümlerden birinin tüm popülasyonu ele geçirmesi durumuna yol açabilir.

Başlangıç popülasyonu, algoritmanın ilk aşamasında, farklı değerlere sahip genlerden oluşan bireyler ile oluşturulur. Başlangıç popülasyonundaki her birey, *uygunluk* adı verilen bir amaç değerine sahiptir. Çözüm uzayında tarama, uygunluk eğer bir maliyet unsuru ise, en küçüklük; bir kazanç unsuru ise, en büyüklük arayışı olarak yapılır. Bu çalışmada uygunluk değeri, seçilen tahmin değişkenleri ile çalıştırılan RO yöntemi ile yapılan sınıflandırma tahmin hatalarına bağlı olarak hesaplanır. Hata sayısı arttıkça kötü kredilerden kaynaklanan maliyetler de artacağından, en düşük hata değerine sahip bireyin en iyi çözüm olarak ele alınması gerekir. Başlangıç popülasyonunun oluşturulma şekli de da çalışmadan çalışmaya farklılık göstermekle birlikte amaç, çözüm uzayı taramasını genişletmek amacıyla olabildiğince farklı bireyler oluşturmak olarak genelleştirilebilir.

Başlangıç popülasyonu ve sonrasında oluşturulacak yeni bireyler için dikkat edilmesi gereken, oluşturulan bireylerin *uygulanabilir* olması gerektiğidir. Bazı problemler için, kullanılan kodlamanın her kombinasyonunun çözümü anlamlı olarak temsil etmesi zor ya da imkânsız olabilir. Böyle durumlarda genel yaklaşım, ya bireyin uygunluk değerine yüksek bir ceza puanı ekleyerek aday çözümler arasında girmesini engellemek ya da direkt olarak popülasyon dışarısına çıkartarak yerine yeni bir birey oluşturmaktır. Bu çalışmada bireyler, rassal olarak atanmış "0" ve "1" değerleri ile oluşturulmuş ve başlangıç popülasyonunu meydana getirmişlerdir. Bu oluşturma yöntemi ile uygulanabilir olmayan bireylerle karşılaşma imkânı yoktur.

### 2.2.3.3. Seçim

GA’da seçim, yüksek uygunluk değerlerine sahip bireylerin kullanılarak yeni bireyler oluşturulması ve böylece mevcut iyi çözümlerin geliştirilmesini amaçlamaktadır. Ancak direkt olarak iyi bireylerin seçimi, olası daha iyi çözümlerin araştırılmasında problem yaratabilmektedir. Bu şekilde yapılan bir seçim, oluşturulacak yeni popülasyonun hızla *yakınsamasına*, sonuçta algoritmanın performansını düşüren *prematüre yakınsama* olgusunun ortaya çıkmasına sebep olacaktır. Prematüre yakınsama, yüksek uygunluğa sahip bir bireyin, birkaç algoritma çevrimi sonrası tüm popülasyonu ele geçirmesidir. Bir yandan mevcut iyi bireylerin yeni popülasyon oluşumunda katkısını yüksek düzeyde tutmak, diğer yandan da mevcut kötü bireylerden ileride ortaya çıkartılabilecek olası daha iyi çözümleri araştırmak gerekmektedir. Ebeveyn olarak seçilecek bireylerin seçimine yönelik yazında farklı seçenekler bulunsa da temelde üç farklı stratejiden bahsetmek mümkündür:

Rulet Çemberi Seçimi: Bu yönteme göre tüm bireyler, uygunluk oranları değerlerine bağlı olarak seçilmektedir. Buna göre, bir bireyin seçilme olasılığı:

$$(4) \quad p(s_i) = \frac{f(s_i)}{\sum_{k=1}^n f(s_k)},$$

olacaktır. Burada  $s_i$ ,  $i$ . bireyi;  $f(s_i)$ ,  $i$ . bireyin uygunluk değeri ve  $p(s_i)$ ,  $i$ . bireyin seçilme olasılığını göstermektedir. Kısaca bir bireyin seçilme olasılığı, bireyin uygunluk değerinin, tüm bireylerin uygunluk değerleri toplamına oranıdır. (4) numaralı denklem, amacın maksimizasyon olduğu durumda geçerlidir. Bu çalışmada olduğu gibi minimizasyon amaçlı problemlerde olasılıklar, uygunluk değerlerinin tersi alınarak oluşturulabilir.

Bu yöntem, her ne kadar iyi uygunluk değerlerine sahip bireylerin seçimini sağlıyor olsa da uygunluk değeri diğerlerinden çok daha iyi olan bireyin popülasyonu hızla ele geçirmesine sebep olacaktır ve çözüm uzayının taranma performansını düşürecektir.

Doğrusal Sıralı Seçim: Doğrusal sıralı seçiminde, rulet çemberi seçiminde karşılaşılan ele geçirme probleminin önüne geçmek amacıyla bireyler uygunluk

değerlerine göre sıralanırlar. Sıralama, maksimizasyon amaçlı çalışmalarda artan, minimizasyon amaçlı çalışmalarda ise azalan şekilde yapılır ve seçim olasılıkları sıralarına bağlı olarak oluşturulur. Yönteme göre her bir bireyin seçim olasılığı:

$$(5) \quad p(s_i) = \frac{r(s_i)}{\sum_{k=1}^n r(s_k)}$$

olacaktır. Burada  $r(s_i)$ ,  $i$ . bireyin sırasını göstermektedir. Örneğin, sıralamanın  $s_3, s_4, s_2, s_1$  olduğu durumda, en yüksek seçim olasılığına sahip birey  $s_1$  (4/10), ikinci yüksek seçim olasılığına sahip birey  $s_2$  (3/10), üçüncü yüksek seçim olasılığına sahip birey  $s_4$  (2/10) ve en düşük seçim olasılığına sahip birey  $s_3$  (1/10) olacaktır.

Turnuva Seçimi: Bu strateji ile öncelikle, popülasyondaki bireyler arasından belirli sayıda birey rassal olarak seçilir ve bir küme oluşturulur. Sonrasında ise küme içerisinde en iyi uygunluk değerine sahip olanı çaprazlama yapılacak ebeveynlerden birisi olarak seçilir. Her yeni ebeveyn seçimi, birbirinden bağımsız seçim kümelerinde yapılan değerlendirmeler ile gerçekleştirilir. Tez çalışmasında turnuva seçimi yöntemi kullanılmıştır.

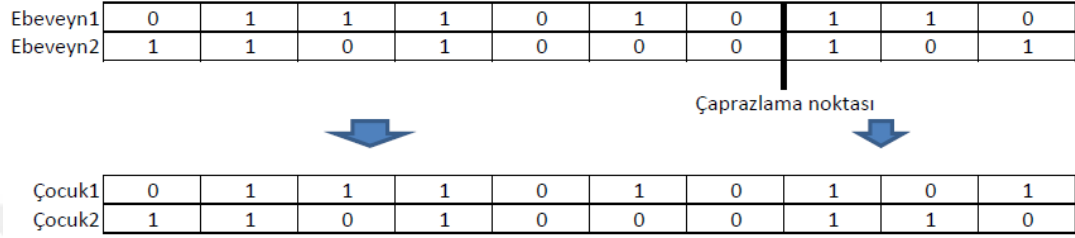
Yeni bireylerin (*çocuk*) oluşturulması için mevcut popülasyondan iki adet birey seçilir. Aynı bireyin ikinci ebeveyn olarak seçilmesi, ilk seçim ile belirlenmiş ebeveynin aday bireyler arasından çıkartılmasıyla engellenmektedir. Belirlenen iki ebeveyn ile iki çocuk oluşturulur ve çocuklar aday birey popülasyonuna dahil edilir. Aday birey popülasyonu, bir sonraki nesli oluşturacak bireylerin seçiminde kullanılmak üzere eski popülasyon ile birlikte değerlendirilecektir. İzleyen iki bölümde, ebeveyn seçimi sonrası çocukların oluşturulmasında kullanılan yöntemler tanıtılacaktır.

#### 2.2.3.4. Yeniden Oluşturma

*Yeniden oluşturma*, seçilmiş bireylerin genlerinin birleştirilmesi işlemidir. GA terminolojisinde buna *çaprazlama* denilmektedir. Yeniden oluşturma mantığı, iyi bireylerin sahip olduğu genlerin birleştirilerek daha iyi bireyler oluşturulabilmesi üzerine kuruludur.

Ebeveynlerin hangi genlerinin çaprazlanacağı ve çaprazlama işleminin gerçekleşme biçimi birey kodlamasına göre farklılık göstermektedir. Bu çalışmada, *tek nokta* çaprazlama kullanılmıştır. Tek nokta çaprazlama örneği, Şekil 9'da gösterilmiştir.

**Şekil 9:** Tek nokta çaprazlama ile yeni bireylerin oluşturulması



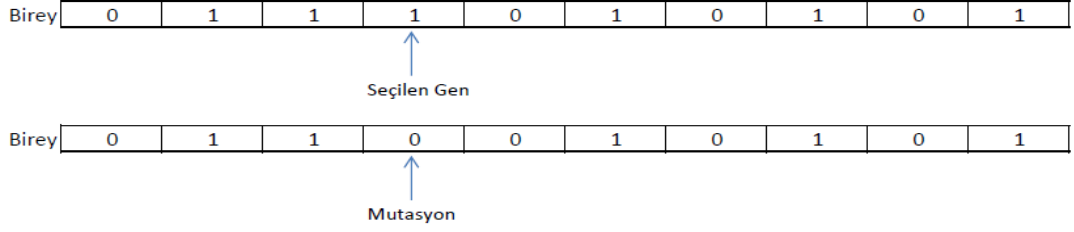
Bu şekilde gerçekleştirilen çaprazlamada, ebeveynlerin çaprazlama noktasına kadar olan genleri aynen alınır ve kalan genler, diğer ebeveynin çaprazlama noktası sonrası kısmından oluşturulur. Bu şekilde, iki ebeveyninden iki çocuk üretilmektedir.

### 2.2.3.5. Mutasyon

*Mutasyon*, yeniden oluşturma sonrası elde edilen çocuklardaki genlerde, belirli bir olasılık dahilinde değişim yaşanmasıdır. Mutasyona uğrayacak olan çocuk bireyde rassal olarak seçilen gen, birey kodlamasına bağlı olarak yeni bir değer alır. Bu çalışmada ikili kodlanmış bireyler bulunduğundan, seçilen gen değeri, karşıt değerine dönüştürülerek mutasyon gerçekleştirilir (0 ise 1; 1 ise 0). Örnek bir mutasyon işlemi Şekil 10'da gösterilmiştir.

Mutasyon bileşeni, GA'nın yerel en iyi çözüme takılmasını engelleyerek çözüm uzayının farklı noktalarının araştırılmasını sağlamaktadır.

**Şekil 10:** İkili kodlanmış bireylerde mutasyon



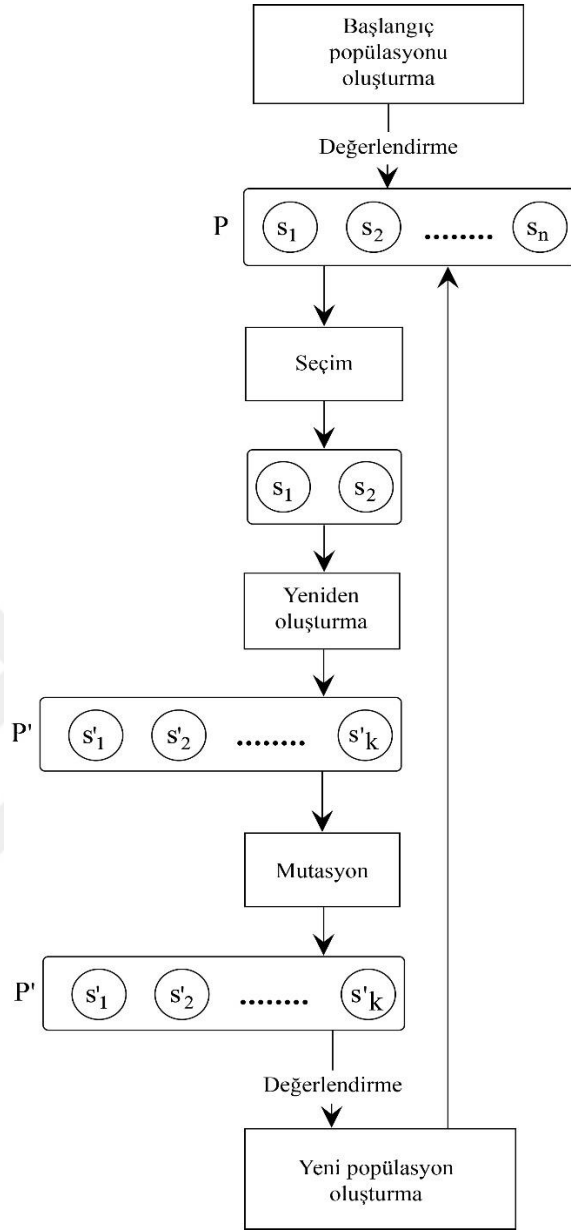
### 2.2.3.6. Yeni Popülasyon Oluşturma ve Genetik Algoritma Adımları

Bir sonraki nesli oluşturacak popülasyonun oluşturulmasında hem eski hem de aday popülasyondan bireyler kullanılmaktadır. Eldeki popülasyonlardan ne şekilde aktarım yapılacağı, *seçkincilik* ile belirlenir. Seçkincilik, eski popülasyondaki iyi bireylerden kaçının yeni popülasyona dahil edileceğinin, kaçının aday popülasyondaki iyi bireylerden oluşturulacağını belirler. Önceden bahsedilen prematüre yakınsama durumu ile karşılaşmamak ve çözüm uzayı araştırmasında farklı bireyleri kullanarak daha etkili olmak amacıyla seçkincilik oranı iyi belirlenmelidir. Bu çalışmada eski ve aday popülasyondan yarı oranda seçim yapılmaktadır. GA'nın tek çevrim akışı Şekil 11'de gösterilmiştir.

Şimdiye kadar bahsedilen işlemler, önceden belirlenmiş bir koşul sağlanana kadar tekrar edilmektedir. Bu tekrarların her biri ile yeni bir *nesil* oluşturulmaktadır. Toplam oluşturulacak nesil sayısı, GA için bir sonlandırma parametresidir. Bu çalışmada da belirli sayıda nesil oluşturulduktan sonra GA çözüm uzayı taraması sona erdirilmektedir. GA çalışma adımlarını şu şekilde sıralamak mümkündür:

- Adım 1: Başlangıç popülasyonu oluştur,
- Adım 2: Bireylerin uygunluk değerlerini hesapla,
- Adım 3: Yeniden oluşturma/çaprazlama,
- Adım 4: Mutasyon,
- Adım 5: Çocukların uygunluk değerlerini hesapla,
- Adım 6: Yeni popülasyon oluştur,
- Adım 7: Maksimum nesil sayısına ulaşılmamışsa Adım 3'e dön,
- Adım 8: En iyi uygunluk değerine sahip bireyi nihai çözüm olarak seç.

**Şekil 11:** Genetik algoritma tek çevrim akışı



#### 2.2.4. Tavlama Benzetimi

TB'nin çalışma mekanizmasının esin kaynağı, katıların çok yüksek derecelere kadar ısıtılarak önce kristal yapılarının bilinçli olarak bozulması ve ardından kontrollü ve çok yavaş şekilde soğutulması işlemidir. Tavlama adı verilen bu işlem ile malzemenin, mümkün olabilecek en uygun kristal yapıya ulaşması sağlanabilir ve böylece kristal kusurları yok edilebilir (Gendreau ve Potvin, 2010, s. 2).



Tavlama süreci benzetimi, optimizasyon problemlerine uygulanabilecek şekilde birbirinden bağımsız olarak Černý (1985) ve Kirkpatrick, Gelatt ve Vecchi (1983) tarafından geliştirilmiştir. Analojiye göre başlangıçta, yüksek bir harekete geçme değeri ile tarama sürecine başlayan sistem her aşamada çözüm uzayını taramakta ve bunu, aşama değerini önceden belirlenmiş bir nihai değere kadar azalttığı her değerlendirme adımında tekrarlamaktadır. TB, birazdan açıklanacak olan mekanizmaları sayesinde, iyi olmayan çözümleri de (belirli bir olasılık dahilinde) kabul ederek, optimizasyon problemlerinde sıklıkla karşılaşılan *yemel en iyi* çözüme takılma sorununu aşabilme gücüne sahiptir.

Sezgisel bir optimizasyon stratejisi olarak TB dizaynı, *Metropolis algoritması* (Metropolis, Rosenbluth, Rosenbluth, Teller ve Teller, 1953) üzerine kurulmuştur. Algoritma, çözüm uzayı taramasını, optimizasyon problemlerinde karşılığı çözümlerin uygunluk değerleri olan sistem enerji seviyeleri arası geçişlerle sağlar. Mevcut enerji halinden (mevcut çözüm uygunluk değeri) daha düşük enerji haline (daha düşük enerji hali daha iyi çözüme karşılık gelmektedir) geçiş koşulsuz kabul edilirken, daha yüksek enerji halinin kabulü, *Metropolis kabul kriterine* bağlı olarak belirlenmektedir. Enerji kavramı, maksimizasyon problemlerinde uygunluk değerleriyle zıt yönlü bir ilişki içerisindedir; düşük enerji hali daha yüksek uygunluk değerine karşılık gelirken yüksek enerji hali tam tersini ifade eder.

Kriteri daha genel bir formda tanımlamak için enerji yerine maliyet kavramı kullanılır. Bir çözümün maliyeti, önceden de belirtildiği gibi maksimizasyon problemlerinde çözümün uygunluk değeri ile negatif yönlü ilişki içerisindedir, bu çalışmada olduğu gibi minimizasyon problemlerinde pozitif yönlü ilişki içerisindedir. Kriteria göre:

$$(6) \quad P(\Delta C) = e^{-\Delta C/T},$$

formülü ile hesaplanan olasılık değeri, kabulün gerçekleşip gerçekleşmeyeceğini belirler. Burada  $\Delta C$ , iki hal çözüm arasındaki maliyet farkı ( $C_{\text{yeni}} - C_{\text{eski}}$ ) olarak ifade edilir. Maliyetler, minimizasyon probleminde uygunluk değerlerinin kendisine karşılık gelir. Ancak uygunluk değeri olarak kullanılan test hata oranları 1'den küçük olduğu için ve hesaplanan farklarının  $e$  tabanlı üstel fonksiyon sonucunda çok küçük

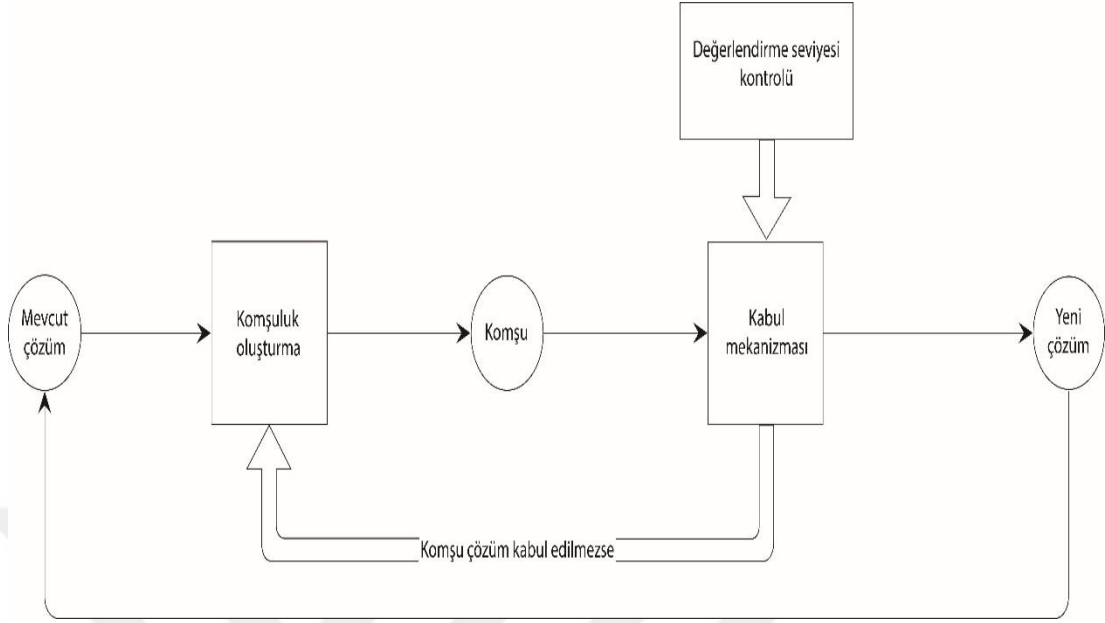
değişiklikler yaratacak olması sebebiyle mevcut çözüm uygunluk değerleri:  $1/f(S_{mevcut})$ ; komşuluk çözüm uygunluk değerleri:  $1/f(S_{komşuluk})$  işlemi ile değiştirilmiştir. Yeni uygunluk değerleri ile problem amacı maksimizasyon haline geldiğinden (daha iyi çözümün test hata oranında sağladığı azalma,  $1/f(S_{komşuluk})$  ve  $1/f(S_{mevcut})$  işlemleri ile artışa karşılık gelir; örneğin  $0.10 < 0.16$  iken  $1/0.10 > 1/0.16$  olur), bir çözümün maliyeti  $C_{\text{çözüm}}$  yeni uygunluk değerinin “-1” ile çarpımıyla elde edilir. T mevcut halin harekete geçme değeridir. İzleyen bölümlerde TB bileşenleri ve kavramları tanıtılacaktır.

#### 2.2.4.1. Çözüm Gösterimi ve Komşuluk Oluşturma

TB’nde kullanılan çözüm de GA’ya benzer şekilde, Bölüm 2.2.3.1’de anlatıldığı ve Şekil 8’de gösterildiği şekliyle oluşturulmaktadır. TB, GA’da olduğu gibi bir birey popülasyonu ile değil, tek bir çözüm gösterimi üzerinden çözüm uzayı taraması yapmaktadır. GA’da mevcut çözümler arasından ebeveynlerin seçimi, çaprazlama ve mutasyon ile oluşturulan aday bireyler yerine TB, mevcut çözüm üzerinden *komşu çözüm* oluşturarak bir ya da daha fazla aday çözüm oluşturur (*komşuluk*).

Komşu çözümlerin nasıl oluşturulacağı ve komşuluğun büyüklüğü, probleme, kodlama biçimine ve dizayn eden kişiye göre değişiklik gösterebilir. Temelde iki tip komşuluk oluşturma yönteminden bahsedilebilir: “Değiştir ve Seç” ve “Değiştir ve Kabul için Kontrol et”. İlk yöntemde, mevcut çözüm gösterimi üzerinden belirli sayıda aday çözüm oluşturularak bir komşuluk kümesi oluşturulur ve çözümler arasından en iyi komşu çözüm seçilir. İkincisinde ise tek bir komşu oluşturularak kabul edilip edilmeyeceğine karar verilir kabul edilirse komşu çözüm mevcut çözümün yerini alır, edilmezse mevcut çözümden yeni bir komşu çözüm oluşturulur.

**Şekil 12:** Tavlama benzetimi çözüm oluşturma süreci



Kaynak: Zäpfel ve Braune, 2010, s. 117.

Bu çalışmada ikinci yöntem benimsenmiştir. Komşuluk oluşturulurken, mevcut çözümün çözüm değerlerinden (GA terminolojisinde gen olarak ifade edilen hücre değerleri) bir tanesi rassal olarak seçilir ve GA'ın mutasyon işlemine benzer şekilde karşıt değeri ile değiştirilir. Mevcut ve komşu çözümün uygunluk değerleri karşılaştırıldıktan sonra komşuluk, iyileştirme sağlıyorsa ya da sağlamıyor ancak Metropolis kabul kriterini sağlıyorsa komşu çözüm kabul edilir; sağlamıyorsa mevcut çözüm üzerinden yeniden komşuluk oluşturulur (Şekil 12).

#### 2.2.4.2. Simülasyon Bileşenleri ve Adımları

TB, bir *harekete geçme-  $T_0$* - değeri ile başlar ve tavlama boyunca azaltılarak *nihai değerlendirme değerine* ulaşınca son bulur. Her sıcaklık seviyesinde, mevcut çözüm üzerinden komşuluk oluşturulur ve *seviye uzunluğu* kadar sayıda komşuluğa geçiş yapılması beklenir.

Harekete geçme değeri, çözüm uzayının yeterli şekilde taranması amacıyla ve özellikle başlangıç seviyelerinde, kötü çözümlerin yüksek olasılıklarla kabul edilmesini sağlayacak şekilde belirlenmelidir. Her sistem değerlendirme seviyesinin

sahip olduğu mükemmel denge halinin (o seviyedeki en iyi çözümün) bilinmesi zaman kısıtları açısından neredeyse imkânsız olduğundan, mevcut çözümden kaç defa komşuluk oluşturulması istendiği bilgisi, yani seviye uzunluğu da simülasyon performansını etkileyen bir diğer önemli bileşendir.

Sistem değerlendirme seviyesi, her seviyede seviye uzunluğu defa komşuluk oluşturulduktan sonra, belirlenmiş bir programa bağlı olarak ayarlanır/azaltılır. Bu programa *tavlama süreci* adı verilir. Diğer bileşenler gibi, tavlama süreci de simülasyon performansı üzerinde büyük etkiye sahiptir. Temel olarak iki tip tavlama süreci vardır: *Statik* veya *uyarlamalı*. Statik süreçte, sistem sıcaklık değeri çeşitli yöntemler ile belirlenen değer kadar azaltılır. Doğrusal bir statik şemada yeni seviyenin sıcaklığı:

$$(7) \quad T_{k+1} = T_k - c,$$

olacaktır. Burada  $T_k$ ,  $k$ . değerlendirme aşamasındaki değerlendirme değeri,  $c$ , sabit bir azaltma değeridir. Geometrik şemada ise yeni seviye:

$$(8) \quad T_{k+1} = T_k * \alpha,$$

şeklinde belirlenir.  $\alpha$ , sabit bir çarpım faktörüdür ve 0 ile 1 arası değer alır. Bu yöntem ile düşük sıcaklık seviyelerinde daha uzun süreli tarama yapılması sağlanır.

Uyarlamalı süreçte ise yeni sıcaklık değeri, sistemin mevcut bilgilerine bağlı olarak belirlenebilir. Örneğin, amaç değerinin istatistiksel özelliklerini dikkate alan bir süreç dizayn edilebilir. Bu çalışmada, *geometrik* şemalı statik bir süreç kullanılmıştır.

Tavlama, daha fazla iyileştirmenin etkin olarak elde edilemediği seviyelerde durdurulmalıdır. Bu seviyeler algoritmanın nihai değeri olarak atanır.

TB adımlarını şu şekilde sıralamak mümkündür:

- Adım 1: Başlangıç çözümü oluştur, en iyi çözüm olarak belirle,
- Adım 2: Harekete geçme değeri belirle, sistem değerlendirme seviyesi olarak ata,
- Adım 3: Seviye uzunluğu belirle,

- Adım 4: Sistem değerlendirme seviyesi, nihai değerlendirme değerinden küçük ise Adım 14'e git,
- Adım 5: Seviye uzunluğu sayacını sıfırla,
- Adım 6: Seviye uzunluğu sayacı seviye uzunluğuna eşit ise Adım 12'e git,
- Adım 7: Seviye uzunluğu sayacını 1 artır,
- Adım 8: Yeni bir komşu oluştur,
- Adım 9: Komşu çözüm ile mevcut çözüm uygunluk değeri farkını hesapla,
- Adım 10: Komşu çözüm uygunluğu daha iyi ise komşu çözüme hareket et, Adım 6'ya dön,
- Adım 11: Metropolis kabul kriteri koşulları sağlanırsa komşu çözüme hareket et, Adım 6'ya dön,
- Adım 12: Mevcut çözüm en iyi çözümden daha iyiyse en iyi çözümü güncelle,
- Adım 13: Sistem değerlendirme değerini güncelle, Adım 4'e dön,
- Adım 14: En iyi çözümü nihai çözüm olarak seç.

### 2.2.5. Birleştirilmiş Algoritmalar- Rastgele Ormanlar Dizayını

Önceki bölümlerde tanıtılmış olan iki meta-sezgisel algoritma, GA ve TB, birçok optimizasyon çalışmasında başarıyla kullanılmaktadır. Bununla birlikte, her iki yaklaşım, çözüm uzayını tarama anlamında çeşitli eksikliklere sahiptir.

GA'lar çözüm uzayını etkin ve hızlı şekilde tarayarak en iyi ya da en iyiye yakın çözümün bulunmasını sağlayabilmektedirler. Bu anlamda GA'lar TB'ne göre daha başarılı gözükmektedirler. Ancak, özellikle yerel en iyi çözümlere takılma problemi ile karşılaştığında yetersiz kalan GA'ları, TB'nin Metropolis algoritma stratejisi ile geliştirmek mümkündür. Böylelikle, *yoğunlaşma ve çeşitlenme* arasında dengeli bir ilişki kurularak genel çözüm uzayı tarama performansı artırılabilir. Yoğunlaşma, çözüm uzayında yapılan taramanın, en iyi çözümün bulunduğu düşünülen dar bir bölgeye odaklanması iken çeşitlenme, çözüm uzayının daha geniş alanlarının uzak geçişler ile taranmasıdır. Yerel en iyiye takılma problemi yaşanması

muhtemel bir GA süreci sonucu elde edilen çözüm, TB yardımı ile yoğunlaşma bölgesinden farklı noktalara da geçiş yapabilmeye şansı elde eder.

Bu çalışmada, iki yöntem birlikte kullanılmış ve yeni algoritmaya BA adı verilmiştir. Buna göre, çözüm uzayının taranması GA ile başlamakta ve ulaşılan çözüm, TB'ne aktarılmaktadır. Aktarılmış çözümü başlangıç çözümü olarak kullanan TB ile varsa yerel en iyiye takılma problemi aşmakta ve çözüm uzayının başka noktalarına geçiş yapılabilir. Çözüm uzayında atlama yapma mekanizması mutasyon bileşeni aracılığı ile GA'da da bulunmakla birlikte Metropolis seçim stratejisi mutasyona göre çok başarılı performans sergilemektedir.

Bölüm 2.2.4.1'de tanımlanan çözüm gösterimi kullanılarak, Bölüm 2.1.2'de tanımlanan RO yönteminin, hangi değişkenler ile öğrenme gerçekleştireceği ve sınıflandırma tahmini yapacağı belirlenmektedir. Değişken seçimine özellikle örneklem büyüklüğünün az, tahmin değişkeni sayısının fazla olduğu *yüksek boyut* varlığı durumunda sıklıkla başvurulmaktadır. Yüksek boyut probleminin var olmadığı durumlarda da tahmin edilen değişkenle ilgisi olmayan ya da farklı kombinasyonlarda daha iyi öğrenme olanaklarının var olduğu zamanlarda anlamlı tahmin performans artışları sağlanabilmektedir. Bağımsız değişkenlerin farklı kombinasyonlarda tahmin değişkeni için ne derece bilgi sağlıyor olduğu bilgisi baştan bilinemediğinden ve ilişkiler genellikle *örtük* olduğundan, en iyi kombinasyon ancak tüm seçenekler denendikten sonra tespit edilebilir. Bu tür problemler, boyut arttıkça hızla uzayan çözüm sürelerine maruz kalmaktadır. Yalnızca bilgisayarların işlem gücünü artırmak bu tür durumlarda yeterli olamamakta, çözüme yönelik olarak yeni yaklaşımların önerilmesi/kullanılması zorunlu hale gelmektedir. Hesaplama karmaşıklığı adı verilen bu olguya ve karmaşıklık sınıflarına ilişkin bilgi bir sonraki verilmiştir.

BA- RO yöntemine ilişkin adımlar şu şekildedir:

- Adım 1: GA başlangıç popülasyonu oluştur,
- Adım 2: Bireylerin uygunluk değerlerini hesapla,
- Adım 3: Yeniden oluşturma/çaprazlama,
- Adım 4: Mutasyon,
- Adım 5: Çocukların uygunluk değerlerini hesapla,
- Adım 6: Yeni popülasyon oluştur,

- Adım 7: Maksimum nesil sayısına ulaşılmamışsa Adım 3'e dön,
- Adım 8: En iyi uygunluk değerine sahip bireyi çözüm olarak seç,
- Adım 9: Adım 8'deki çözümü başlangıç çözümü olarak seç, en iyi çözümü olarak belirle,
- Adım 10: Harekete geçme değeri belirle, sistem değerlendirme seviyesi olarak ata,
- Adım 11: Seviye uzunluğu belirle,
- Adım 12: Sistem değerlendirme seviyesi, nihai değerlendirme değerinden küçük ise Adım 20'ye git,
- Adım 13: Seviye uzunluğu sayacı seviye uzunluğuna eşit ise Adım 19'a git,
- Adım 14: Yeni bir komşu oluştur,
- Adım 15: Komşu çözüm ile mevcut çözüm uygunluk değeri farkını hesapla,
- Adım 16: Komşu çözüm uygunluğu daha iyi ise komşu çözüme hareket et, seviye uzunluğu sayacını 1 artır, Adım 18'e git,
- Adım 17: Metropolis kabul kriteri koşulları sağlanırsa komşu çözüme hareket et, seviye uzunluğu sayacını 1 artır, Adım 18'e git,
- Adım 18: Mevcut çözüm en iyi çözümden daha iyiyse en iyi çözümü güncelle, Adım 13'e dön
- Adım 19: Sistem değerlendirme seviyesini güncelle, seviye uzunluğu sayacını sıfırla, Adım 12'ye dön,
- Adım 20: En iyi çözümü nihai çözüm olarak seç.

Özetle, meta-sezgisel algoritmaların oluşturduğu çözümlere bağlı olarak seçilmiş olan değişkenler ile veri kümesini öğrenen RO, kredi sınıflandırma tahmin hatalarını algoritmalara uygunluk değeri olarak aktarmaktadır. En iyi uygunluk değerini veren çözüm, en iyi değişken kombinasyonu olarak seçilmektedir.

## 2.2.6. Genetik Algoritma ve Tavlama Benzetimi Parametre Değer ve Seçimleri

Meta-sezgisel algoritmaların bileşen parametre seçimleri, çözüm uzayı tarama performansları üzerinde direkt etkili olmaktadır ancak her çalışmada en iyi performansı veren ortak değerler/seçimlerden bahsetmek mümkün olamamaktadır. Değerler büyük ölçüde problemlere ve algoritma dizaynlarına bağlıdır ve deney tasarımları çerçevesinde denemelerle belirlenmektedir. Yine de genel kabul gören bazı yaklaşımlar mevcuttur.

**Tablo 5:** Genetik algoritma parametreleri

Parametre	Değer/Seçim
Popülasyon büyüklüğü	30
Nesil sayısı	50
Seçim Stratejisi	Turnuva- popülasyonun 4'te 1'i
Çaprazlama şekli	Tek nokta
Mutasyon oranı	%20
Yerine koyma şeması	Seçkincilik- eski popülasyonun yarısı

GA'lar için popülasyon, erken yakınmaya sebep vermeyecek ancak işlem süresini uzatmayacak büyüklükte olmalıdır. Yine benzer şekilde nesil sayısı, işlem süresi dikkate alınarak çözüm uzayının yeterince taranmasına imkân verecek şekilde belirlenmelidir. Prematüre yakınsama problemine takılmamak ve çözüm uzayının daha geniş alanlarının taranmasının sağlanması için mutasyon oranının da iyi belirlenmesi şarttır. Bu çalışmada GA'ya ait ve yapılan denemeler sonucu iyi performans verdiği tespit edilen, Bölüm 2.2.3'te bahsi geçen operatörlerin aldığı değerlere ve seçimlere ilişkin bilgiler Tablo 5'te sunulmuştur.

TB parametre seçimlerinin öneminden Bölüm 2.2.4.2'de bahsedilmiştir. GA parametre seçimlerine benzer şekilde denemeler ile tespit edilen TB parametre değerlerine ilişkin bilgiler de Tablo 6'da bulunmaktadır.



**Tablo 6:** Tavlama benzetimi parametreleri

Parametre	Değer/Seçim
Harekete geçme değeri	300
Seviye uzunluğu	5
Tavlama şeması	0.9 oranı ile Geometrik
Nihai değerlendirme değeri	7

### 2.3. PERFORMANS DEĞERLENDİRME ÖLÇÜTLERİ

Çalışmada kullanılan yöntemlerin kredi risk tahmin performanslarını değerlendirmek amacıyla *karışıklık matrisine* başvurulmuştur. Karışıklık matrisi, salt tahmin hatasının değil, ele alınan probleme göre önemi değişen farklı hata cinslerinin de hesaplanmasını sağlayan bir araçtır. Kredi sınıflandırması için örnek bir karışıklık matrisi Tablo 7’de gösterilmiştir.

**Tablo 7:** Kredi sınıflandırması için karışıklık matrisi

Gözlemlerin tahmin edilen sınıfları	Gözlemlerin gerçek sınıfları	
	KÖTÜ (pozitif)	İYİ (negatif)
KÖTÜ (pozitif)	Doğru pozitif (DP)	Yanlış pozitif (YP)
İYİ (negatif)	Yanlış negatif (YN)	Doğru negatif (DN)

Burada, doğru pozitif (DP), başarılı şekilde sınıflandırılmış “kötü” başvuruların; doğru negatif (DN), başarılı şekilde sınıflandırılmış “iyi” başvuruların; yanlış pozitif (YP), hatalı sınıflandırılmış “kötü” başvuruların; yanlış negatif (YN), hatalı sınıflandırılmış “iyi” başvuruların sayısıdır. Kredi risk çalışmalarında amaç olarak genellikle, temerrüde düşme ihtimali olan başvuruların tespiti daha ön planda olmaktadır. Bunun sebebi, gerçek sınıfı “iyi” olan bir başvurunun reddedilmesinin, gerçek sınıfı “kötü” olan bir başvurunun kabul edilmesine göre etkisinin daha az olmasıdır. Başarısız bir kredi sürecinin sonucunda ortaya çıkacak zararın boyutu, tersi durumdaki kazanca göre hem mikro hem de makro ölçekte çok daha büyüktür.

Kredi tahmin performanslarının farklı ölçütler ile değerlendirilmesi için üç farklı kesinlik değerine başvurulabilir. *Tahmin kesinliği*, bu çalışmada da temel performans göstergesi olarak kullanılan ve tüm gözlemler arasında doğru şekilde sınıflandırılan kredi başvurularının oranıdır. Tahmin kesinliği değeri:

$$(9) \quad \text{Tahmin kesinliđi} = \frac{DP+DN}{DP+DN+YP+YN},$$

formülü ile hesaplanır. *TİP-I kesinlik*, başarılı şekilde tahmin edilen “kötü” kredilerin oranıdır ve *hassasiyet* olarak da bilinir. TİP-I kesinlik:

$$(10) \quad \text{TİP - I kesinlik} = \frac{DP}{DP+YN},$$

formülü ile hesaplanır. Son gösterge *TİP-II kesinlik*, bir diđer adıyla *özgüllük*, başarılı şekilde tahmin edilen “iyi” kredilerin oranıdır ve:

$$(11) \quad \text{TİP - II kesinlik} = \frac{DN}{YP+DN},$$

formülü ile hesaplanır.

Sınıflandırma için kullanılan öğrenen yöntemlerin performanslarının değerlendirilmesinde dikkat edilmesi gereken, sistemin eğitildiđi veri kümesini, bir diđer deyişle eğitim kümesindeki gözlemler için tahmin başarısının temel hedef olmadığıdır. Bu tür çalışmalarda amaç, öğrenen yöntemin genelleme yeteneđine sahip olması, böylece eğitim sırasında kullanılmayan gözlemleri başarılı şekilde sınıflandırabilmesidir.

Yöntemin, *eđitim kümesi* gözlemlerini tahminindeki hata oranına *eđitim hatası*, yeni gözlemleri, yani *test kümesini* tahminindeki hata oranına ise *test hatası* adı verilir. Genel olarak tüm öğrenen yöntemlerin eğitim hataları, test hatalarından daha düşük olmaktadır. Eğitim kümesindeki rassal hatalar ve gürültüler tarafından fazlasıyla etkilenen öğrenen yöntemin genelleme yeteneđinin kaybolduđu bu durumlara *aşırı-uyum gösterme* adı verilir. Bu sorunun önüne geçmek için kullanılabilircek çeşitli yöntemler bulunmaktadır.

Bu yöntemler ile tüm veri kümesi, eğitim ve test kümelerine bölünür ve öğrenen yöntemin eğitimi için eğitim kümesi kullanılırken, performans değerlendirmesi için eğitim kümesi ile oluşturulmuş modelin test kümesi üzerindeki tahminlerine başvurulur. Eğitim ve test kümesi belirlemeye ve böylece aşırı uyum

gösterme problemini aşmaya ilişkin yöntemler temelde iki gruba ayrılmaktadır. Bunlardan ilkinde, k-katlamalı çapraz-onaylama adı verilmektedir. Buna göre tüm veri kümesi eşit boyutta k farklı ve birbirinden bağımsız test kümesine bölünür ve her k test grubu sınıflandırma tahmini, geriye kalan grupların sınıflandırıcı ile eğitimi sonrası gerçekleştirilir. Bu şekilde tahmin çalışması yapılan tüm gruplara ilişkin sonuçlar toplandıktan sonra da bunların ortalaması alınarak sınıflandırmanın performansı ortaya çıkartılır.

Diğer yöntem ise onaylama kümesi adı verilen, k-katlamalı çapraz-onaylamaya benzer şekilde test gruplarının oluşturulduğu, ancak bunun tüm veri kümesini kapsayacak şekilde yapılması yerine kümeler dahil olacak gözlemlerin rassal olarak belirlendiği yöntemdir. Kimi çalışmalarda, aynı gözlemlerin birden fazla kez eğitim kümesine dahil edilmesi de söz konusu olabilmektedir.

Her iki yöntemde de amaç, bağımsız olarak incelenen veri kümeleri üzerindeki performansın daha iyi ortaya konmasıdır. İyi bir sınıflandırma, test hatasını minimize eden, aşırı-uyum sorununun var olmadığı sınıflandırmadır. Bu amaçla, çalışmada kullanılan kredi veri kümesi, eğitim kümesi ve test kümesi olarak ikiye ayrılmış ve eğitim kümesi ile eğitimi sonrası yöntemin performans değerlendirmeleri, test kümesini tahmin, TİP-I ve TİP-II kesinlikleri dikkate alınarak yapılmıştır.

Tüm öğrenen yöntemlerin temel hedefi, eldeki gözlemlerin kullanılarak bağımsız değişkenler ile tahmin değişkeni arasındaki gerçek ilişkinin ortaya çıkartılmasını sağlamaktır. Ancak pratikte bu ilişki hiçbir zaman tam anlamıyla ortaya çıkartılamamakta, veri kümesindeki gözlemlerde tahmin edilemeyen sapmalardan - *gürültü*- ya da modelleme yetersizliklerine bağlı hatalardan kaynaklanan bozulmalar olmaktadır. Eğitimin hangi sayıda gözlem ile daha başarılı olacağı ve gerçek ilişkinin tam olarak hangi gözlemler ile sağlanabileceği de bilinemediğinden, yöntem performansı değerlendirmesinde, tüm veri kümesi içerisinde farklı sayılarda rassal seçilmiş gözlemlerin kullanılması değerlendirmenin daha sağlıklı yapılabilmesi açısından yerinde olacaktır. Bu amaçla veri kümesi 100, 200, 300, 400, 500 adet gözlemden oluşan test kümelerine ayrılmış ve geriye kalan gözlemler yöntemin eğitimi için kullanılmıştır. Her test kümesi büyüklüğü için de 20 farklı ve rassal olarak oluşturulmuş gözlem grupları kullanılmıştır. Performanslar, her test küme büyüklüğünde oluşturulmuş gözlem gruplarının tahmin, TİP-I ve TİP-II başarılarının

ortalaması alınarak değerlendirilmiştir. Bu şekilde farklı gözlemler içeren eğitim ve test kümesi büyüklerinin performans üzerinde nasıl bir etkisi olduğu da araştırılmış olacaktır. Farklı eğitim kümeleri kullanılmasıyla ortaya çıkacak performans farklılaşmaları da, tahmin hatalarının standart sapmalarının hesaplanmasıyla ortaya konmuştur.

Üçüncü bölümde, analiz edilen veri kümesi tanıtılmış ve analiz yönteminin çalışma akışına ilişkin şemalar sunulmuştur. Bulguların aktarılmasından sonra Bölüm sonuçların tartışılması ile son bulmaktadır.



## ÜÇÜNCÜ BÖLÜM

### ÖRNEK KREDİ DEĞERLENDİRME VERİ KÜMESİ ÜZERİNDE BİR İNCELEME

Bu bölümde ilk olarak, kullanılan veri kümesi ve incelemenin amacından bahsedilecektir. Ardından, RO yöntemi ile elde edilmiş olan bulgular ve sonrasında da BA-RO bulguları, Bölüm 2.3'te bahsedilmiş olan performans değerlendirme ölçütleri çerçevesinde sunulacaktır. İki yöntem performans karşılaştırılması, Bölüm 3.5'te yapılacak ve Bölüm 3.6'da sonuçlar, yazındaki çalışmalar ile karşılaştırılacaktır.

#### 3.1. İNCELEMENİN AMACI

Önerilen yöntemin başarısının doğru değerlendirilebilmesi açısından, kredi risk değerlendirme çalışmalarında başvurulan bir veri kümesinin analiz edilmesi gerekmektedir. Birçok çalışmada kullanılan ve karşılaştırma amaçlı olarak kendine sıklıkla yer bulan bir veri seti olarak, “UCI Machine Learning Repository” veri tabanı üzerinden erişilebilen “Alman Tüketici Kredi Veri Kümesi” kullanılmıştır (Lichman, 2013). Kredi değerlendirme çalışmalarında kullanılan veri kümelerine ilişkin kapsamlı bilgi için Lessmann, Baesens, Seow ve Thomas (2015)'in çalışmasına başvurulabilir.

1000 gözlemden oluşan veri seti içerisinde hem nitel hem de sayısal olarak 20 tahmin değişkeni- bağımsız değişken bulunmaktadır (Tablo 8). Çalışmanın kalan kısmında değişkenlere ait numaralar kullanılacaktır. Değişkenlere ilişkin açıklamalar EK 2'de verilmiştir. Ayrıca her gözlemin, kredi temerrüdüne düşüp düşmediğine ilişkin olarak “iyi” ve “kötü” değerleri alabilen bir sınıflandırma değişkeni vardır. Sınıflandırma yöntemlerinin amacı, veri kümesindeki gözlemleri, sunulan değişkenler ile değerlendirerek doğru şekilde sınıflandırmaktır. Değişkenlere ilişkin betimleyici istatistikler EK 3'te verilmiştir.

**Tablo 8:** Çalışmada kullanılan veri kümesine ait değişken bilgileri ve kategorik özellikleri

No	Değişken	Kategori
De.1	Vadesiz mevduat hesap durumu	Nitel
De.2	Kredi vadesi -Aylık	Sayısal
De.3	Kredi geçmişi	Nitel
De.4	Amaç	Nitel
De.5	Kredi miktarı	Sayısal
De.6	Tasarruf hesabı/ tahviller	Nitel
De.7	Çalışma süresi	Nitel
De.8	Harcanabilir gelirin yüzdesi olarak taksit oranı	Sayısal
De.9	Medeni durum ve cinsiyet	Nitel
De.10	Ortak başvuru/ kefiller	Nitel
De.11	Mevcut ikametgâh süresi	Sayısal
Değ12	Sahip olunan mülkler	Nitel
De.13	Yaş	Sayısal
De.14	Ödeme planı	Nitel
De.15	Barınma	Nitel
De.16	Bankada bulunan kredi sayısı	Sayısal
De.17	İş durumu	Nitel
De.18	Bakmakla yükümlü olduğu kişi sayısı	Sayısal
De.19	Telefon	Nitel
De.20	Yabancı çalışan	Nitel

Günümüzde, kuşkusuz olarak, daha fazla ve farklı değişken içeren kredi değerlendirme süreçlerinden bahsetmek mümkündür. Her şeyden önce, mevcut ekonomik koşulları ele almayan bir analiz yetersiz kalacaktır. Ancak, değerlendirme yöntemi önerisi çalışmalarında nesnel ve doğru bir karşılaştırma yapmak ihtiyacı da ortadadır. Yazında aynı veri kümesi üzerinde yapılan çalışmalarla (Abellán ve Castellano, 2017; Ala'Raj ve Abbod, 2016a, 2016b; Florez-Lopez ve Ramon-Jeronimo, 2015; C.-L. Huang ve diğerleri, 2007; Oreski ve Oreski, 2014; Wang, Ma, Huang ve Xu, 2012; Yu, Yao, Wang ve Lai, 2011; Zhang, Leung, ve diğerleri, 2008; Zhao ve diğerleri, 2015) performansı değerlendirilen yöntem, güncel piyasa ve ekonomi verileri ve başvurular ile ilgili elde edilen yeni bilgiler ile de kolaylıkla uygulanma şansına sahiptir.

### 3.2. ÖNERİLEN ANALİZ YÖNTEMİNİN GENEL AKIŞI

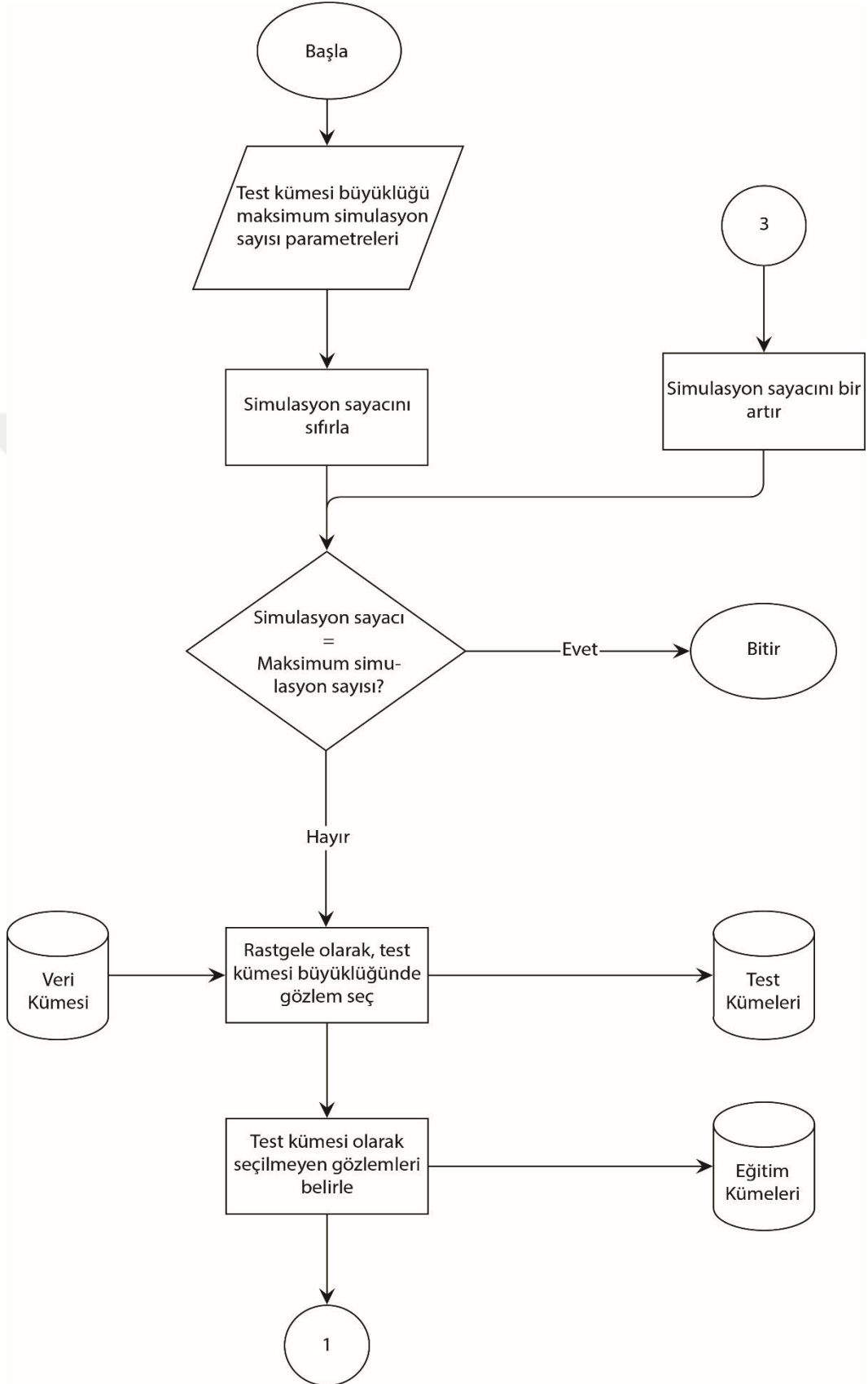
Çalışmada önerilen veri kümesi analiz süreci, Şekil 13'te gösterildiği gibi eğitim ve test kümelerinin oluşturulmasıyla başlamaktadır. Gerek RO yöntemi ile ağaçların oluşturulma süreci, gerekse BA ile çözüm üretme süreci rassal süreçler olduğundan, model duyarlılığının anlaşılması için farklı eğitim kümeleri kullanılan 20 farklı simülasyon tasarlanmıştır. Her simülasyonda hem RO yöntemi hem de RO-BA yöntemleri ayrı ayrı çalıştırılmıştır; böylelikle RO yönteminin tek başına kullanımı ile tez çalışmasında önerilen yöntem anlamlı şekilde karşılaştırılabilmektedir.

İkinci aşamada, BA önerisinin GA kısmı çalıştırılmaktadır. Şekil 14'te görüldüğü gibi, test ve eğitim kümelerinin belirlenmesinden sonra başlangıç popülasyonu oluşturulmakta, RO yöntemi ile popülasyondaki bireyler değerlendirildikten sonra GA'nın "Ebeveyn Seçimi", "Çaprazlama", "Mutasyon" ve "Aday Popülasyon Oluşturma" aşamaları gerçekleştirilmektedir. Aday popülasyondaki bireyler de RO yöntemi ile değerlendirildikten sonra hem aday hem de mevcut bireylerin performanslarına bağlı olarak yeni popülasyon oluşturulmaktadır. GA, önceden belirlenen nesil sayısına ulaşıncaya kadar devam etmektedir.

Şekil 15'te gösterilmiş olan üçüncü aşama, GA ile belirlenmiş en iyi çözüm gösterimi TB yöntemine aktarılmasıyla başlamaktadır. TB, mevcut çözüm üzerinden komşuluk oluşturmakta, RO yöntemi ile analiz edilen komşu çözümün kabulü performans ölçütü veya Metropolis kriteri ile değerlendirilmekte ve süreç, nihai değerlendirme değerine ulaşıncaya kadar devam etmektedir. TB süreci sonunda, BA yaklaşımı ile elde edilmiş ve RO analiz performansını en iyi düzeyine çıkarmış olan çözüm, en iyi çözüm olarak kabul edilmektedir.

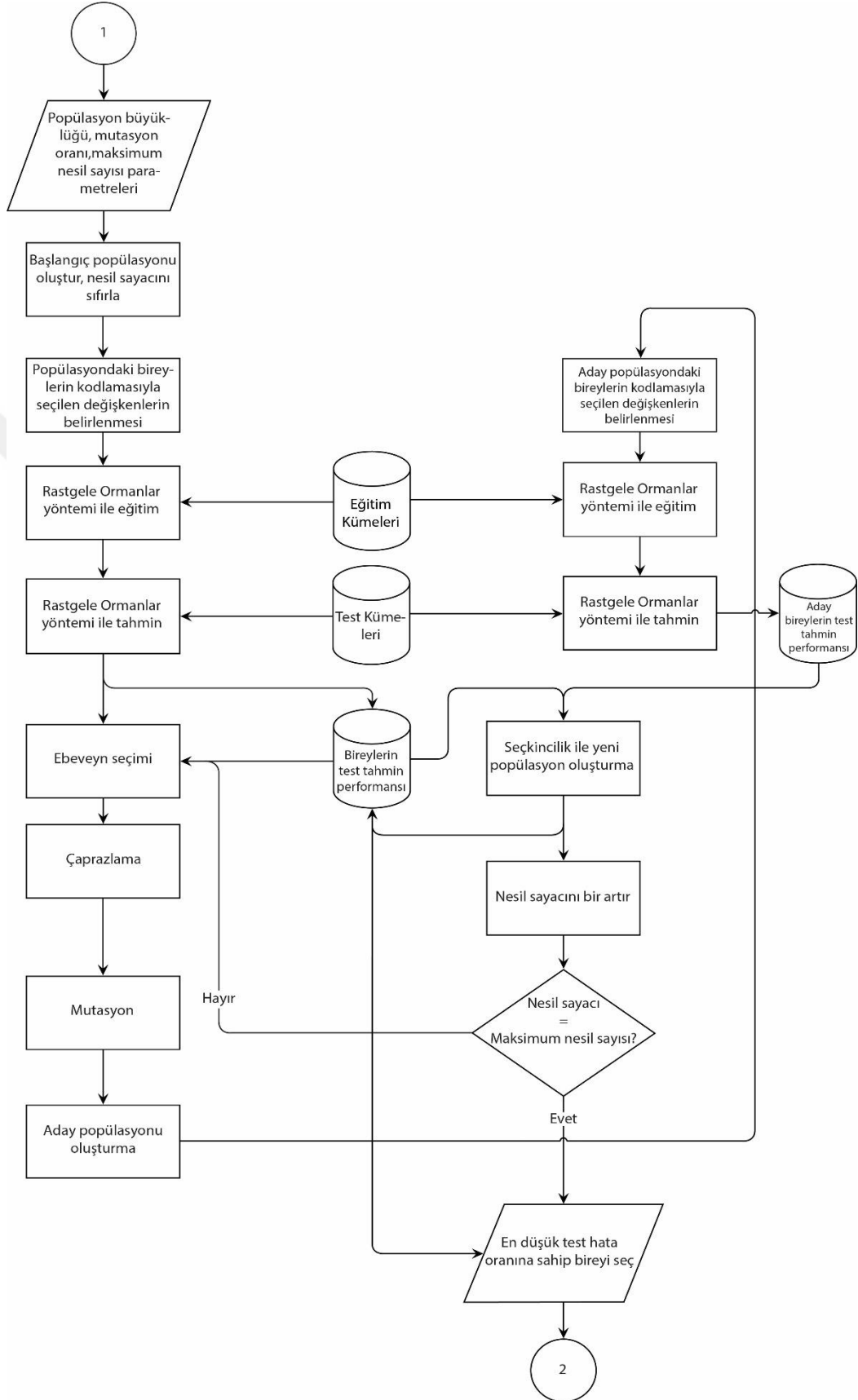
Simülasyonun sonlanmasıyla birinci aşamaya dönülmekte ve yeni simülasyon çalıştırılmaya başlanmaktadır. Bu aşamalar, 100, 200, 300, 400 ve 500 test kümesi büyüklükleri için ayrı ayrı tekrarlanmaktadır.

Şekil 13: Veri kümesinin analiz aşamaları: birinci aşama

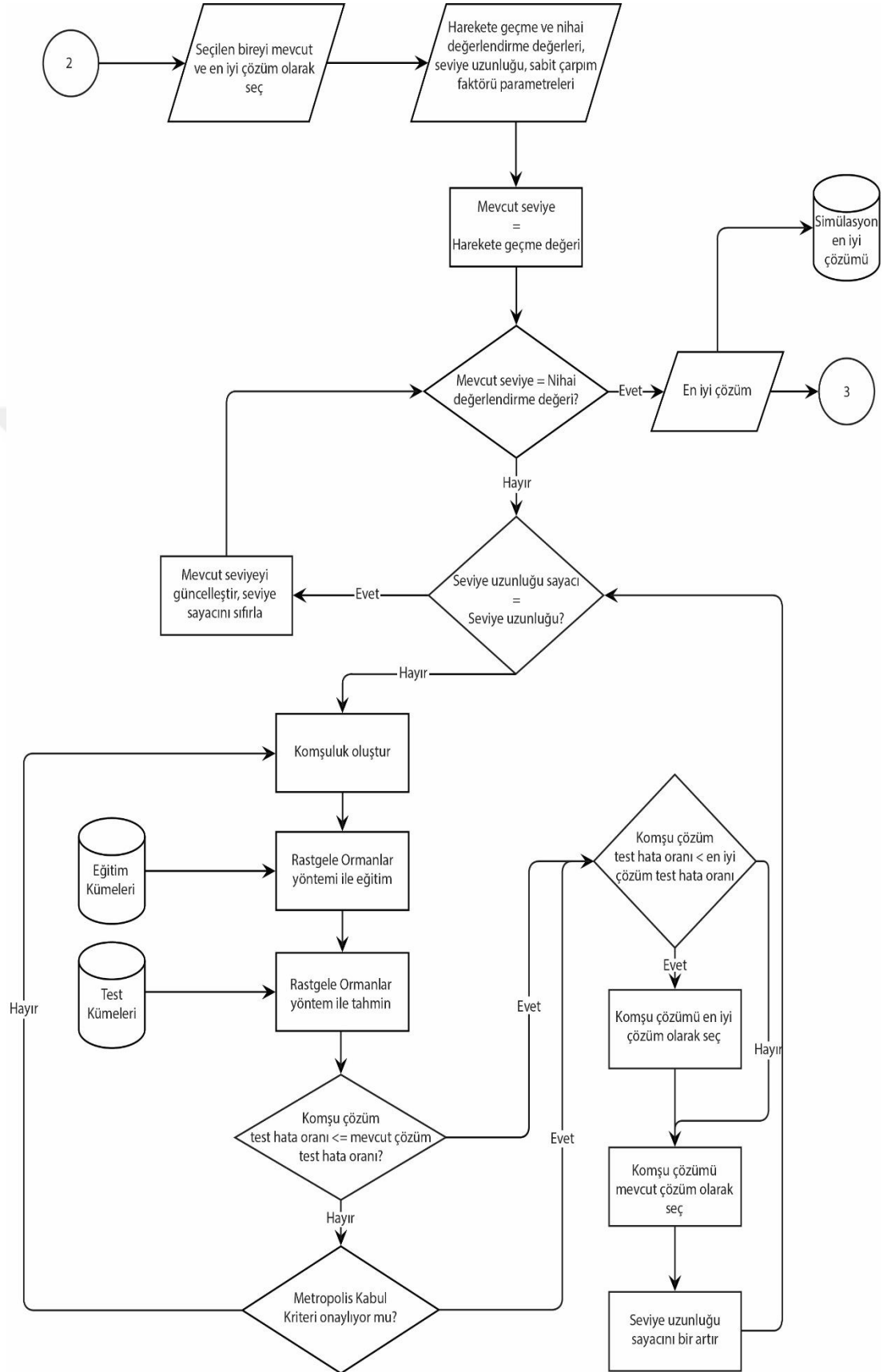




Şekil 14: Veri kümesinin analiz aşamaları: ikinci aşama



Şekil 15: Veri kümesinin analiz aşamaları: üçüncü aşama



### 3.3. RASTGELE ORMANLAR BULGULARI

RO yöntemi bulgularına ilişkin performans ölçütleri, Tablo 9’da gösterilmiştir. Değerler, beş farklı test grubu büyüklüğü için gözlem gruplarının tahmin hata ortalamalarıdır. Önceden de belirtildiği gibi ortalama alınmasının sebebi, her test grubu büyüklüğü için 20 farklı gözlem kümesi oluşturulmuş olmasıdır. Böylece, yöntemin farklı gözlemler ile eğitilmesi durumunda yaşanacak performans değişkenliğine ilişkin de fikir sahibi olmak mümkün olacaktır.

İlk dikkat çeken, farklı test kümesi büyüklüklerinin tahmin performansında ciddi değişikliklere sebep olmadığıdır. Tahmin kesinliği, 100, 200, 300, 400, 500 test kümesi büyüklükleri için sırasıyla %76.1, %76.4, %77, %76.2 ve %74.9 olmuştur. Bunda, RO yönteminin tahmin yeteneğinin farklı eğitim kümesi büyüklüklerinden fazla etkilenmemesi, aşırı-uyum gösterme probleminin ortaya çıkmaması etkili görülmektedir.

Standart sapma değerleri incelendiğinde ise, test kümesi büyüklüğü artığında farklı gözlem kümelerinin tahmin hatalarındaki sapmanın azaldığı tespit edilmiştir. Az sayıda eğitim gözlemi ile daha fazla sayıda test gözleminin tahmin edilmesinin nispeten güvenilir sonuçlar verdiği söylenebilir. Eğitim kümesi tahminleri, tüm test kümesi büyüklükleri için %100 başarılı olmuştur. Bu açıdan yöntemin eğitim kümesini başarılı şekilde öğrendiğini söylemek mümkündür.

TİP-I kesinlikler oldukça düşük çıkmıştır. TİP-II kesinlikler ile karşılaştırıldığında TİP-I kesinlikte gözlenen sorun daha da belirgin olmaktadır. Göstergeler RO yönteminin “kötü” kredi başvurularını doğru tahmin etmekte, “iyi” kredi başvurularını tahmin etmeye göre daha başarısız olduğuna işaret etmektedir. Buna, veri kümesindeki gözlemlerin “iyi” ve “kötü” gözlemler bakımından dengesiz bir dağılıma sahip olması neden olabilir. Bu farklılıkların, değerlendirmenin amacına ve olası sonuçlarına göre göz önüne alınması yerinde olacaktır. Son olarak, yöntemin ortalama işlem süreleri incelendiğinde, beklendiği gibi, eğitim kümesi büyüklüğü azaldıkça süresinin de düştüğü görülmektedir. Bu farklar, pratikte çok önemli gözükmemektedir.

**Tablo 9:** RO yöntemi ile kredi riski tahmin sınıflandırma bulguları

Test Kümesi Büyüklüğü	100			
	Test Tahmin Hatası	Test Tahmin Hatası Standart Sapma	Eğitim Tahmin Hatası	Eğitim Tahmin Hatası Standart Sapma
	0.239	0.046	0.000	0.000
TİP-I Kesinliği	0.413			
TİP-II Kesinliği	0.914			
İşlem Süresi (sn.)	5.531			
Test Kümesi Büyüklüğü	200			
	Test Tahmin Hatası	Test Tahmin Hatası Standart Sapma	Eğitim Tahmin Hatası	Eğitim Tahmin Hatası Standart Sapma
	0.237	0.028	0.000	0.000
TİP-I Kesinliği	0.400			
TİP-II Kesinliği	0.925			
İşlem Süresi (sn.)	4.941			
Test Kümesi Büyüklüğü	300			
	Test Tahmin Hatası	Test Tahmin Hatası Standart Sapma	Eğitim Tahmin Hatası	Eğitim Tahmin Hatası Standart Sapma
	0.230	0.018	0.000	0.000
TİP-I Kesinliği	0.432			
TİP-II Kesinliği	0.910			
İşlem Süresi (sn.)	4.256			
Test Kümesi Büyüklüğü	400			
	Test Tahmin Hatası	Test Tahmin Hatası Standart Sapma	Eğitim Tahmin Hatası	Eğitim Tahmin Hatası Standart Sapma
	0.238	0.014	0.000	0.000
TİP-I Kesinliği	0.387			
TİP-II Kesinliği	0.925			
İşlem Süresi (sn.)	3.600			
Test Kümesi Büyüklüğü	500			
	Test Tahmin Hatası	Test Tahmin Hatası Standart Sapma	Eğitim Tahmin Hatası	Eğitim Tahmin Hatası Standart Sapma
	0.251	0.014	0.000	0.000
TİP-I Kesinliği	0.374			
TİP-II Kesinliği	0.914			
İşlem Süresi (sn.)	2.950			

### 3.4. BA-RO BULGULARI

Önceki bölümde RO yöntemi ile analiz edilen eğitim kümeleri birebir olarak, BO-RO yönteminde kullanılmıştır. Geliştirilmiş eğitim bulgularına ilişkin performans ölçütleri, Tablo 10’da gösterilmiştir. Değerler, beş farklı test grubu büyüklüğü için gözlem grupları tahmin hata ortalamalarıdır. Ayrıca, her test kümesi büyüklüğünde kullanılan 20 farklı gözlem gruplarının tahminlerinin minimum, maksimum ve ortanca tahmin hata değerleri de verilmiştir. Tabloda eğitim kümesi tahmin hatalarına yer verilmemiştir.

Meta-sezgisel algoritmalar ile geliştirilmiş RO, tüm test kümesi büyüklüklerinde artan bir performans göstermiştir (Tahmin başarıları 100, 200, 300, 400 ve 500 test kümesi büyüklükleri için sırasıyla: %83.5, %80.6, %79.6, %79.1, %78.3). Tahmin hatalarının standart sapmaları, yalnızca RO yöntemi ile elde edilen analiz bulgularında olduğu gibi test kümesi büyüklüğü artıkça azalmaktadır. Minimum test tahmin hataları arasında en iyi performans yine 100 test kümesi büyüklüğünde gerçekleşmiştir. Tüm test kümesi büyüklüklerinde birbirine yakın maksimum hata oranları ortaya çıkmıştır. Ortalama değerlere yakın ortanca değerler, farklı gruplarının tahmin performanslarının simetrik bir dağılım gösterdiğine işaret etmektedir.

100 gözlem içeren test kümeleri, TİP-I ve TİP-II kesinlikleri için de diğer test kümesi büyüklüklerine göre daha iyi performans göstermişlerdir. Burada da test kümesi büyüklüğü artıkça azalan kesinlik değerlerinden bahsetmek mümkündür. TİP-I kesinlikleri, gözlemlerdeki “iyi” ve “kötü” dengesizliğinden etkilenmeye devam etmektedir.

Artan performanslar karşılığında, eğitim sürecinin tüm test kümesi büyüklükleri için ciddi şekilde yükseldiği gözlenmektedir. Özellikle daha çok gözlemin eğitim için kullanıldığı 100 test kümesi büyüklüğünde eğitim süresi, 500 test kümesi büyüklüğüne göre iki katından fazla (3248,086 sn. daha fazla) olmuştur.

**Tablo 10:** BA-RO yöntemi ile kredi riski tahmin sınıflandırma bulguları

Test Kümesi Büyüklüğü	100		
Test Tahmin Hatası	0.165	Test Tahmin Hatası Standart Sapma	0.031
TİP-I Kesinliği	0.583	Min. Test Tahmin Hatası	0.110
TİP-II Kesinliği	0.939	Ortanca Test Tahmin Hatası	0.165
İşlem Süresi (sn.)	6023.259	Maks. Test Tahmin Hatası	0.240
Test Kümesi Büyüklüğü	200		
Test Tahmin Hatası	0.194	Test Tahmin Hatası Standart Sapma	0.019
TİP-I Kesinliği	0.519	Min. Test Tahmin Hatası	0.160
TİP-II Kesinliği	0.926	Ortanca Test Tahmin Hatası	0.190
İşlem Süresi (sn.)	5847.419	Maks. Test Tahmin Hatası	0.230
Test Kümesi Büyüklüğü	300		
Test Tahmin Hatası	0.204	Test Tahmin Hatası Standart Sapma	0.017
TİP-I Kesinliği	0.504	Min. Test Tahmin Hatası	0.170
TİP-II Kesinliği	0.921	Ortanca Test Tahmin Hatası	0.205
İşlem Süresi (sn.)	4736.921	Maks. Test Tahmin Hatası	0.240
Test Kümesi Büyüklüğü	400		
Test Tahmin Hatası	0.209	Test Tahmin Hatası Standart Sapma	0.015
TİP-I Kesinliği	0.487	Min. Test Tahmin Hatası	0.175
TİP-II Kesinliği	0.919	Ortanca Test Tahmin Hatası	0.211
İşlem Süresi (sn.)	3257.742	Maks. Test Tahmin Hatası	0.243
Test Kümesi Büyüklüğü	500		
Test Tahmin Hatası	0.217	Test Tahmin Hatası Standart Sapma	0.015
TİP-I Kesinliği	0.490	Min. Test Tahmin Hatası	0.190
TİP-II Kesinliği	0.908	Ortanca Test Tahmin Hatası	0.216
İşlem Süresi (sn.)	2775.173	Maks. Test Tahmin Hatası	0.242

### 3.5. ANALİZ BULGULARININ KARŞILAŞTIRILMASI

Geliştirilen yöntemin sağladığı performans artışının istatistiksel olarak anlamlı olup olmadığını anlamak için, parametrik olmayan Wilcoxon testi kullanılmıştır. Test, aynı örnek gruplarının karşılaştırıldığı analiz sonuçlarının karşılaştırılmasında kullanılmaktadır ve özellikle küçük örnek büyüklüklerinde bağımlı örneklem t-testine göre daha doğru sonuç verebilme kapasitesine sahiptir. Test sonuçları Tablo 11’de gösterilmiştir. Düşük p-değerleri, tüm test kümesi büyüklüklerindeki tüm gözlemler için artışın anlamlı düzeyde olduğunu göstermektedir.

**Tablo 11:** Wilcoxon testi sonuçları

Test Kümesi Büyüklüğü	p-değeri
100	0.00*
200	0.00*
300	0.00*
400	0.00*
500	0.00*

\*%99 güven düzeyinde anlamlılık

RO ile BA-RO yöntemleri karşılaştırıldığında, tahmin performansındaki en büyük artışın 100 gözlem içeren test kümesinde gerçekleştiği (ortalama %7.5 artış) görülmüştür. Test kümesi büyüklüğü arttıkça, performans artış oranı azalmaktadır. Tüm test kümesi büyüklükleri için TİP-I kesinliklerinde artış sağlanmıştır. Artış miktarları 100, 200, 300, 400 ve 500 büyüklükleri için sırasıyla %17, %11.9, %7.2, %10 ve %11.6 olmuştur. En büyük artış, tahmin performansında olduğu gibi 100 test kümesi büyüklüğünde olmuştur. TİP-II kesinliklerde ise yalnızca 200 ve 300 test kümesi büyüklüklerinde (%0.1 ve %0.6) artış gerçekleşmiştir. Diğer test kümesi büyüklüklerinin TİP-II kesinliklerinde azalma olmuştur (100, 300 ve 500 büyüklükleri için sırasıyla %2.5, %1.1, %0.6). Ancak bu artış ve azalışlar düşük düzeyde olduğu için yöntem performansı için gösterge niteliğinde değillerdir. Son olarak, 100 ve 200 test gözlem kümesi büyüklüğü oynaklıklarını, BA-RO yöntemi ile azaltabilmek mümkün olmuştur. 300, 400, 500 gözlem büyüklüklerinde ciddi bir oynaklık değişkenliği yoktur. Yöntem, bu anlamda da geçerli bir performans artışı sağlamaktadır.

### 3.6. TARTIŞMA

Louzada, Ara ve Fernandes (2016), yakın döneme kadar kredi dereceleme üzerine gerçekleştirilen çalışmaların sistematik bir taramasını yapmışlardır. İnceledikleri analiz çalışmalarının yarısına yakınının değerlendirme için yeni bir yöntem önerisinde bulunduğunu, çok küçük bir kısmının değişken seçimi yaptığını tespit etmişlerdir. Bu anlamda tez çalışmasının yeni yöntem önerisi ile genel eğilimi takip ettiği, önerilen yöntem ile değişken seçimi gerçekleştirerek de özgün bir gruba katkı sağladığı söylenebilir. Söz konusu çalışmanın yöntem tercihleri incelemesinde

ise, çalışmada önerilen çeşitli yöntemlerin birleştirilerek performanslarının artırılması amacıyla, hibrit yöntemler ile birlikte diğer çalışmalarda da ön planda olduğu görülmektedir.

Daha önceden de belirtildiği gibi, önerilen yöntemin başarısını değerlendirmek, aynı veri kümesinin kullanıldığı diğer çalışmalar ile karşılaştırma yaparak mümkün olabilecektir. Alman tüketici kredi veri kümesini kullanarak yöntemlerinin performanslarını sunan çalışmalar a bakıldığında, BA-RO yöntemi ile elde edilen sonuçların fazlasıyla tatmin edici olduğu görülmektedir (bkz. Tablo 5, Louzada ve diğerleri, 2016, s. 128). Tabloda verilen çalışmaların genel tahmin performans ortalaması %77.89 olmuştur ve bu çalışmada önerilen yöntem ile ortalama değer %5.61 üzerine çıkmıştır.

Başarısı yüksek çalışmaların metod tercihleri incelendiğinde, hibritleştirilmiş ya da birleştirilmiş yöntemler içerisinde sinir ağlarının ön plana çıktığı görülmektedir. Hsieh (2005) çalışmasında, kümeleme ile sinir ağlarının bir arada kullanıldığı hibrit bir yöntem kullanmış ve %98.5'lik bir başarı oranına ulaşmıştır. Hibritleştirilmiş sinir ağları yöntemleri ile Marcano-Cedeño, Marin-de-la-Barcena, Jimenez-Trillo, Piñuela ve Andina (2011) %84.67, birleştirilmiş lojistik regresyon ve sinir ağı yöntemi ile Lu, Liyan ve Hongwei (2013) %80, genetik algoritmalar ile hibritleştirilmiş sinir ağları ile Oreski ve Oreski (2014) %78.9, yalnızca sinir ağları yöntemleri ile Khashman (2010) %85.9 (bu çalışmanın performans oranı içerisinde eğitim kümesi de kullanıldığı için test kümesi performansına göre yüksek çıkmış olması olasıdır) ve Zhao ve diğerleri (2015) eğitim kümelerinde eşit sayıda sınıfa ait gözlem olmasına yönelik yaklaşım önerileri ile birlikte SA ile %83-%87 arası performans değerleri elde etmişlerdir. Abellán ve Castellano (2017), torbalanmış DVM ile %76.6 başarı elde ederken Ala'Raj ve Abbod, (2016a)'da çeşitli öğrenen yöntemlerin kararları arasında ağırlıklandırmaya dayalı bir dereceleme ve seçim yapılarak %77.8 tahmin başarısına ulaşılmıştır. Yine aynı yazarlar, Alman tüketici veri kümesinde veri filtrelemesi (aykırı gözlemlerin kümeden çıkartılması) ve değişken seçimi sonrası önerdikleri "fikir birliği" yöntemiyle yaptıkları analizde, %79'luk bir tahmin yüzdesine ulaşmışlardır (Ala'Raj ve Abbod, 2016b). Fikir birliği yöntemi ile, RO, KA, SA, DVM ve Bayes sınıflandırma yöntemlerinin ortaklaşa şekilde karar vermeleri sağlanmış ve böylece her yöntemin sınıflandırma gücünden



yararlanmaya çalışılmıştır. Paralel RO yöntemi kullanan Van Sang ve diğerleri (2016) %76.2'lik bir başarı oranına ulaşmışlardır.

Bu çalışmaların haricinde, değişken seçimi yapmalarını sağlayan yöntemler kullanan daha farklı çalışmalar da bulunmaktadır. Destek vektör makineleri ile genetik algoritmalarını hibritleştiren Huang, Chen ve Wang (2007) %77.92'lik başarı oranına ortalama 13.3 değişken ile, kümeleme ve genetik algoritma kullanan Zhang, Leung ve Ye (2008), %77.76'lık başarıya 16 değişken ile, önceden bahsi geçen Oreski ve Oreski (2014) ise söz konusu başarı oranına ortalama 12.36 değişken ile ulaşmışlardır. Jadhav ve diğerleri (2018), hibrit değişken seçimi kullandıkları çalışmalarında en iyi başarı oranını (%82.8), GA ile yaptıkları değişken seçimi sonrası çalıştırılan DVM ile yakalamışlardır.

**Tablo 12:** Farklı test kümesi büyüklüklerinde seçilen değişken sayısı ve genel ortalaması

Test Kümesi Büyüklüğü	Seçilen Değişken Sayısı
100	10.55
200	12.75
300	11.65
400	12.4
500	11.7
Ortalama	11.81

Bu çalışmada önerilen yöntemin değişken seçimlerinin ortalaması, her test kümesi büyüklüğü için Tablo 12'de gösterilmiştir. Farklı test kümesi büyüklükleri arasında ciddi bir farklılaşma yoktur. En az seçim, en iyi performansın elde edildiği 100 test kümesi büyüklüğünde gerçekleşmiştir. Bulgulardan ulaşılan sonuç, gözlemlere ait mevcut tüm değişkenlerin kullanımından ziyade değişken seçimi yapılmasının, performanslar üzerinde geliştirici etkisi olduğudur. Bu anlamda, değişken seçim yöntemleri değişken uzayı boyutunu azaltarak RO başarısını artırmıştır.

Kredi değerlendirmesinde bir diğer önemli konu, doğru seçimin yapılmasında hangi değişkenin daha büyük önem arz ettiğinin anlaşılmasıdır. Bu, özellikle hangi verilerin müşteriden talep edilmesi gerektiğinin ve öznel değerlendirme yapılması gereken durumlarda neye dikkat edilmesi gerektiğinin bilinmesini sağlayacaktır. Bu

aynı zamanda, başvurusu reddedilen bir müşteriye, ret sebebi ile ilgili olarak anlamlı bir gerekçe sunulması açısından da bir gerekliliktir. BA yöntemi ile değişkenlerin seçim oranları, Tablo 13'te gösterilmiştir.

Oranlar, test kümesi büyüklüklerine göre farklılıklar göstermektedir. Örneğin, değişken 3 olarak kodlanan kredi geçmişinin seçim yüzdeleri, 100 ile 500 test kümesi büyüklükleri arasında %30 kadar farklı olmuştur. Diğer değişkenler için de buna benzer ciddi farklılıklar gözlenmektedir. Burada esas üzerinde durulması gereken, öğrenen yöntemlerin sunulan verilerden bilgi çıkarma biçimleridir.

**Tablo 13:** Farklı test kümesi büyüklüklerinde değişkenlerin seçim yüzdesi ve genel ortalamalar

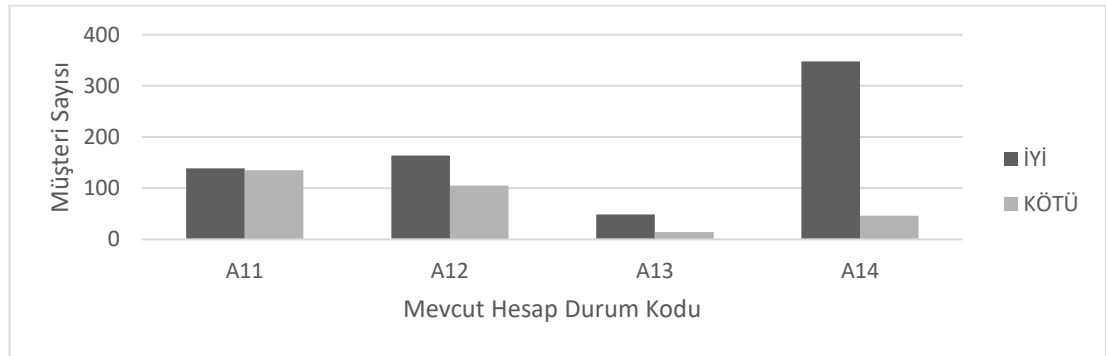
Test Kümesi Büyüklüğü	De. 1	De. 2	De. 3	De. 4	De. 5	De. 6	De. 7	De. 8	De. 9	De. 10
100	%85	%70	%65	%50	%50	%75	%40	%55	%40	%55
200	%90	%80	%75	%60	%80	%70	%80	%65	%45	%60
300	%100	%95	%75	%50	%75	%75	%75	%30	%40	%65
400	%95	%90	%90	%40	%90	%80	%75	%50	%35	%85
500	%100	%95	%95	%45	%80	%95	%65	%35	%45	%75
Ortalama	%94	%86	%80	%49	%75	%79	%67	%47	%41	%68
Test Kümesi Büyüklüğü	De. 11	De. 12	De. 13	De. 14	De. 15	De. 16	De. 17	De. 18	De. 19	De. 20
100	%45	%40	%45	%55	%40	%55	%70	%25	%55	%40
200	%70	%60	%60	%60	%45	%40	%50	%50	%65	%70
300	%60	%65	%65	%35	%25	%60	%30	%35	%60	%50
400	%45	%50	%60	%70	%40	%65	%50	%40	%60	%30
500	%40	%50	%40	%30	%35	%50	%60	%40	%45	%50
Ortalama	%52	%53	%54	%50	%37	%54	%52	%38	%57	%48

Klasik yaklaşımlarda, bir bağımlı değişken ile ilişkisi olduğu düşünülen bağımsız, yani belirleyici değişkenler analize dahil edilir ve tahmin fonksiyonu biçim ve parametrelerinin belirlenmesinin ya da sınıflandırma uzayı ayrıştırmasının bu değişkenlerin tamamı dikkate alınarak yapılması sağlanır. Bu tür analizlerde, ilişkili/ilişkisiz değişkenlerin tespiti ve modelin buna göre kurulması gerekmektedir. Ancak, değişken ekleme/çıkarma, bağımsız değişkenlerin tek başlarına taşıdıkları bilgilerin

yanında diğer değişkenler ile birlikte taşıdıkları bilgilerin ortak etkisinin de ortadan kalkması gibi öğrenmeyi olumsuz etkileyen bir duruma yol açabilir. Bu etkiler, her ne kadar klasik yöntemlerle de bir dereceye kadar tespit edilebiliyor olsa da değişken sayısının fazlaştığı analizlerde etkilerin gözden kaçma olasılığı artacaktır. Artan hesaplama gücü ile kolaylaşan analiz süreçleri ile tezde önerilen yöntem ve benzerleri sayesinde, öznel olarak anlamlı olmayan ya da anlaşılmayan ancak tahmin performansına etki eden söz konusu ilişkileri ortaya çıkartmak mümkündür. Nitekim, değişken seçimi ile artan performanslar da çeşitli çalışmalarda ortaya konmuştur (Bellotti ve Crook, 2009; F. L. Chen ve Li, 2010; C.-L. Huang ve diğerleri, 2007; Koutanaei ve diğerleri, 2015; Oreski ve diğerleri, 2012).

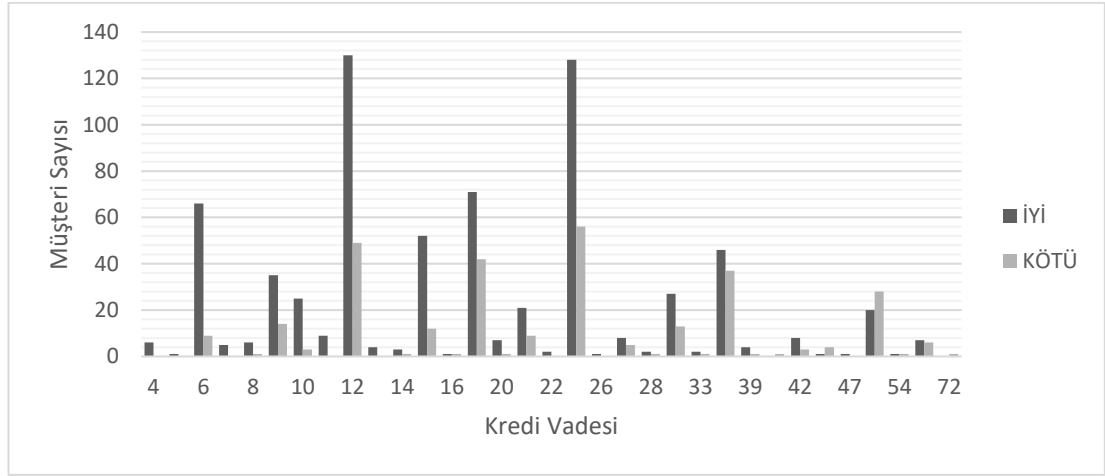
Bu çalışmada da farklı değişken alt grupları ile RO performansını artırmak mümkün olmuştur. Genel ortalamalara bakıldığında, müşterinin, %70'in üzerinde seçim oranlarına sahip vadesiz mevduat hesap durumu (de.1), kredi vadesi (de.2), kredi geçmişi (de.3), kredi miktarı (de.5) ve tasarruf hesabı/tahviller (de.6) bilgileri temerrüt riskinin değerlendirilmesinde ön plandadır. Söz konusu değişken değerler ile “iyi” ve “kötü” müşteri sınıflarının histogramları, sırasıyla Şekil 16, Şekil 17, Şekil 18, Şekil 19, Şekil 20 ve Şekil 21’de gösterilmiştir.

**Şekil 16:** Müşterilerin vadesiz mevduat hesap durumları ve sınıf dağılımları



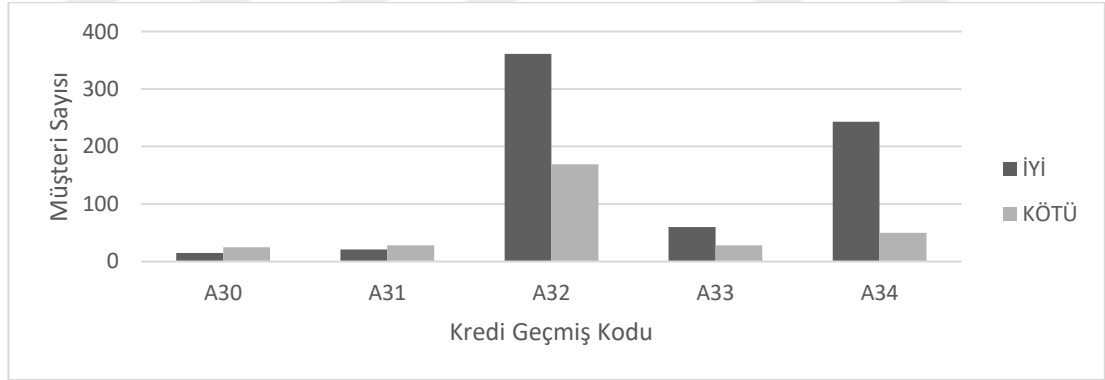
Müşterilerin vadesiz mevduat hesap durumlarına göre temerrüde düşme oranlarına bakıldığında, hesaptaki para miktarı artıkça temerrüt oranının azaldığı, ancak en iyi müşteri grubunun “A14” yani vadesiz hesabı olmayanlar olduğu görülmektedir (“kötü”- “iyi” oranı 0.141). Bu anlamda söz konusu ilişkinin yönünü bu grafik ile açık olarak görmek mümkün değildir.

**Şekil 17:** Müşteri kredilerinin vadesi ve sınıf dağılımları



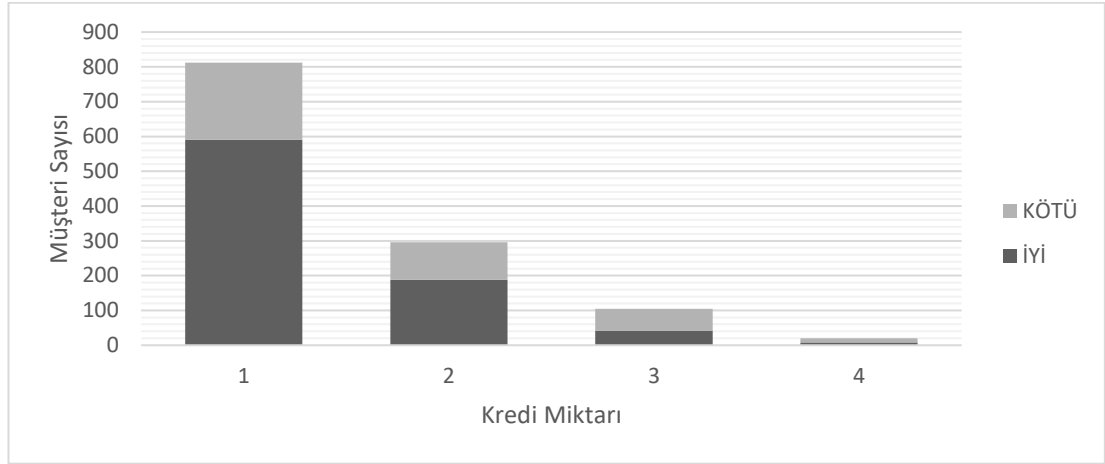
Kredi vadeleri ile temerrüt durumları arasındaki ilişkiye ilişkin kabaca, uzun vadelerde temerrüt miktarının arttığını söylemek mümkündür. Tabii ki bu ilişki doğrusal bir nitelikte görülmemektedir, en kısa ile en uzun vade arasındaki sürelerde temerrüt oranları değişiklik göstermektedir.

**Şekil 18:** Müşterilerin kredi geçmişleri ve sınıf dağılımları



Kredi geçmişleri ile temerrüt ilişkileri ise çok daha ilginç bir karakteristiğe sahiptir. Müşteriler arasından, geçmiş kredilerini ödemelerinde gecikme yaşayan ya da kritik hesap olarak kodlananların, zamanında ödeme yapan ya da daha önce kredi çekmemiş olanlara göre temerrüt oranlarının daha az olduğu görülmektedir. Öznel değerlendirme açısından bu ilişki çok anlamlı görülmemektedir. Ancak RO yönteminin, öznel değerlendirme ile gözden kaçabilen başka bir ilişkiyi yakalamış olması muhtemeldir.

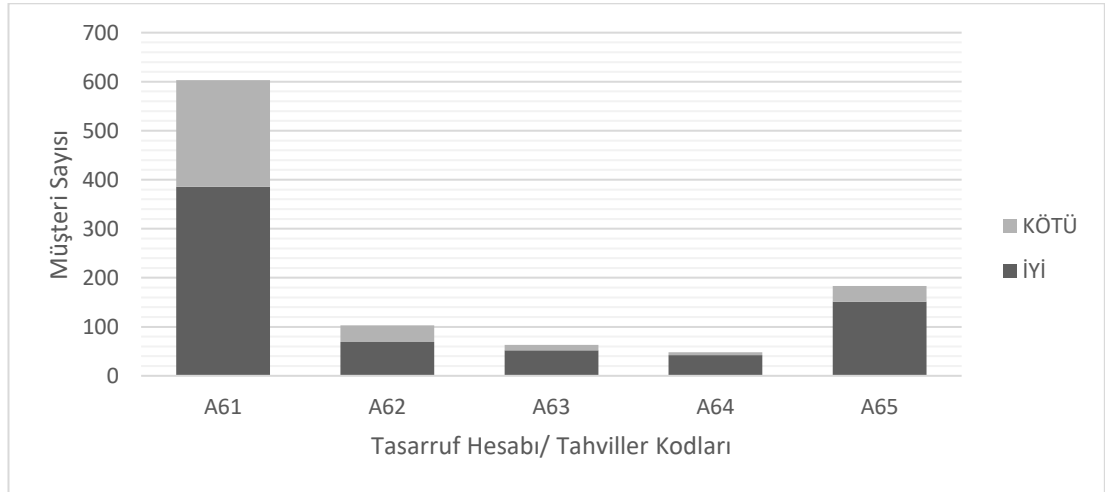
**Şekil 19:** Müşterilerin çektikleri kredi miktarı ve sınıf dağılımları



*Kredi miktarı kodları: (1) <5000 Alman Markı, (2) 5000-9999 Alman Markı, (3) 10000-14999 Alman Markı, (4) ≥15000 Alman Markı.*

Çekilen kredi miktarlarının, temerrüt oranları üzerinde önemli bir etkisi olduğu görülmektedir. Kredi miktarı artıkça, “kötü”- “iyi” oranı sırasıyla 0.376, 0.575, 1.5, 1.5 olmaktadır. Büyük miktarda kredi çeken müşterilerin ödeme güçlüğüne daha çok düştüğünü söylemek mümkündür.

**Şekil 20:** Müşterilerin tasarruf hesap ve tahvil durumları ve sınıf dağılımları



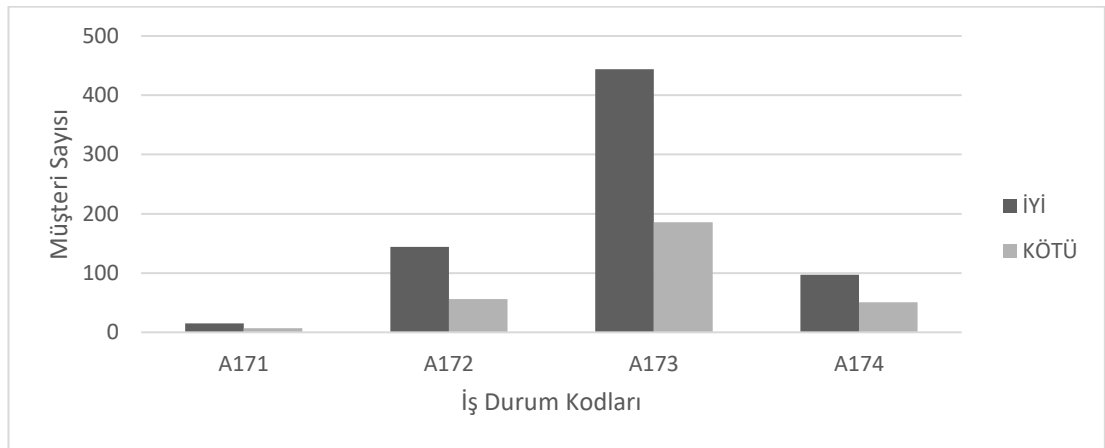
Müşterilerin tasarruf hesapları ve tahvil durumları ile temerrüt ilişkilerine bakılacak olursa, düşük tasarrufa sahip müşterilerin arasında “kötü” sınıfa sahip olanların, diğerlerine göre daha fazla olduğu söylenebilir. Bu açıdan anlamlı görülen

ilişki, hiç hesabı olmayan müşterilerde de “kötü” oranının düşük olmasını tam olarak açıklayamamaktadır.

Görüldüğü gibi, görsel yollarla ya da basit ilişkiler üzerinden müşterilerin sınıflandırılması sağlıklı, çoğu durumda da mümkün olamamaktadır. Bağımlı ve bağımsız değişkenler arasında doğrusal ya da anlaşılır bir doğrusal-olmayan ilişkinin varlığı söz konusu değilse, çok boyutlu çözüm uzaylarında bağımsız değişkenlerin ortak etkilerinin ortaya çıkartılması için öğrenen yöntemlere (ya da benzer işlevi olan diğer yöntemlere) başvurmak bir zorunluluk olmaktadır. Bu anlamda bu çalışmada kullanılan ve geliştirilen yöntemin, bu zorunluğa cevap olarak oldukça kullanışlı olduğunu söylemek mümkündür.

Son olarak, en iyi performansı veren 100 test kümesi büyüklüğü için değişken seçimlerine bakıldığında, %70 üzerinde oranlara sahip vadesiz mevduat hesap durumu (de1), kredi vadesi (de.2), tasarruf hesabı/tahviller (de.6) ve iş durumu (de.17) değişkenleri etkili görülmektedir. Şekil 21’de, iş durumları ile temerrüt sınıfları histogramı verilmiştir. Buna göre, müşterilerin iş durumları ile “kötü”- “iyi” sınıfları oranı arasında ciddi bir farklılaşma yoktur (Sırasıyla 0.467, 0.389, 0.419, 0.526). İş durumunun, diğer değişkenler ile örtük bir etkileşim içerisinde olduğu kuvvetle ihtimaldir.

**Şekil 21:** Müşterilerin iş durumları ve sınıf dağılımları



Tezin sonraki ve son bölümünde, çalışmanın yazına katkısı ve uygulanabilirliği ile ileride gerçekleştirilebilecek çalışmalara yönelik yorumlara yer verilmiştir.

## SONUÇ

Modern finans sisteminin gelişimi boyunca karşılaşılan ve yanlış/eksik değerlendirmelere bağlı ortaya çıkan sorunların bertaraf edilmesinin, sürdürülebilir bir ekonomik ortamın sağlanması açısından önemi ortadadır. Küresel ölçekte güvenilir bir sistem dizaynına ilişkin çalışmalar finansal mimari kavramı altında toplanmaktadır. Finansal mimari, uzun yıllar boyunca tecrübe edilmiş krizlerin sebep ve sonuçlarının irdelenmesi ile sürekli olarak yeniden dizayn edilmektedir. Bu dizayn sürecinde düzenleyici ve denetleyici kurumların sistem içerisinde yer alan aktörler üzerindeki etkisi de sürekli artış göstermektedir. Söz konusu etkiler, 80'li yıllardan beri gündemde olan Basel düzenlemeleri altında toplanmakta ve uygulanmaktadır. Basel I-II ve III düzenlemeleri ile ortaya konan tavsiye ve yönetmeliklerin temel vurgularından biri başarılı risk yönetim süreçlerinin oluşturulmasıdır.

Çeşitlenen finansal araçlar, birleşen ve serbestleşen pazarlar ile geniş bir kredi tabanı oluşturulmuş ve paralel olarak da bankalar ve bireyler için durmaksızın artan bir risk iştahı ortaya çıkmıştır. Genişleyen kredi tabanı sayesinde üretim artmış, artan üretime karşılık tüketim de sunulan kredi imkanları sayesinde canlı tutulmuştur. Özellikle 2007 tarihli küresel finans krizine giden yolun inşasında yüksek rekabet koşulları içerisinde faaliyet gösteren bankaların, geri ödeme yükümlülüklerini yerine getirme potansiyeli düşük tüketicilere kullandığı konut kredilerinin büyük rolü vardır. Artan finansal imkanlar ile birlikte hızla gelişen ve büyüyen ülke ekonomileri, alınan risklerin kontrol edilememesi ve denetlenememesi sebebiyle ciddi yıkımlara da uğrayabilmektedir. Son küresel kriz, bu yıkımın en taze ve büyük örneğidir. Eksikliklerin ve yapılması gerekenlerin büyük ölçüde farkında olan düzenleyici kurumlar, risk iştahlarının sebep olacağı sorunların önüne geçilmesine özel önem vermektedir. Kredi veren kurumların yeterli sermaye gereksinimlerine sahip olması ve böylece olumsuz etkilerin soğurulabilmesi çalışmaların bir ayağını oluşturmakla birlikte, krizlerin ortaya çıkışını engelleyecek mekanizmaların inşa edilmesi çok daha sağlıklı bir sistemin ortaya çıkmasında kilit rol üstlenmektedir.

Makro ölçekte ortaya çıkan problemlerin engellenmesi yolunda mikro ölçekte yaklaşımlar geliştirmek gerekmektedir. Bu kapsamda da kurumların risk yönetimi süreçleri içerisinde gerek karşılaşılan muhtemel kayıpların önlenmesi gerekse

ortaya çıkan kayıplara karşı sistemin işleyiş ve bütünlüğünü tehlikeye atmayacak bir soğurma ortamının oluşturulabilmesi için kredi başvurularının doğru şekilde değerlendirilmeleri odak noktasında durmaktadır. Yöntemsel ve teknolojik gelişmeler, çok daha başarılı değerlendirmelerin yapılması için büyük imkanlar sunmaktadır.

Basel düzenlemeleri çerçevesinde kurumlara çeşitli risk değerlendirme yaklaşımları sunulmaktadır. Bunlar arasında uygulaması en basit ve maliyetsiz olanı, kredi risklerinin sabit risk ağırlıkları ile kategorize edilerek değerlendirdiği standart yaklaşımdır. Standart yaklaşım gelişmiş bir modelleme çalışması gerektirmemektedir ancak kredi başvurularını birkaç tip ağırlık içerisinde değerlendirdiği için fazlasıyla basit ve gerçekçilikten uzak olduğu yönünde yoğun eleştirilere de maruz kalmaktadır. Standart yaklaşım yerine temel ve gelişmiş yaklaşımları içerisinde barındıran içsel değerlendirme yöntemlerinin, yüksek bilişim imkanlarına ulaşılan günümüz koşullarında kullanımı çok daha anlamlı olmaktadır. İçsel değerlendirme yöntemleri ile temerrüt olasılığı, temerrüt halinde kayıp ve temerrüt tutarı gibi risk parametrelerine bağlı olarak her başvuruya ilişkin risk ağırlıkları, başvurulara ait verilerin çeşitli yöntemler aracılığı analizi ile belirlenmektedir. Başvuruların değerlendirilmesinde kullanılacak veri miktarındaki artış ve bu verinin, bu tez çalışmasındaki gibi gelişmiş araç ve yöntemler sayesinde anlamlı bilgiye dönüştürülebilme kapasitesi, içsel değerlendirme çalışmalarının geleceği açısından fazlasıyla umut vericidir.

Bu çalışmada, tüketici kredilerinin temerrüt risklerini yüksek doğruluk oranı ile sınıflandıracak ve böylece bankaların yönetmelikler ve öneriler çerçevesinde içsel olarak gerçekleştirdikleri kredi değerlendirme süreçlerinin başarısını artıracak bir yöntem önerisi sunulmuştur. Yöntemin performansının yazındaki diğer çalışmalar ile anlamlı şekilde karşılaştırılabilmesi amacıyla da Alman tüketici veri kümesi analiz edilmiştir. Önerilen analiz yöntemi, bir makine öğrenme yöntemi olan RO ile birleştirilmiş GA ve TB meta-sezgisel algoritmalarının hibritleştirilmesi ile oluşturulmuştur. BA olarak anılan birleştirme işlemi, GA ve TB'nin ayrı ayrı sahip olduğu güçlü yönlerin etkisinin artırılmasını ve zayıf yönlerin etkisinin azaltılmasını sağlamıştır.

DR, LR, ÇAA gibi öğrenen yöntemleri uzun zamandan beri risk değerlendirme çalışmalarında kendilerine yer bulmaktadır. Ancak bu yöntemler son otuz yıllık süreçte bilgi işleme kapasitelerinde ve istatistiksel/makine öğrenim yöntemlerinde



yaşanan gelişmelere bağlı olarak sık kullanım özelliklerini kaybetmişlerdir. Bunun başlıca sebebi, artan bilgi işleme kapasiteleri sayesinde incelenen veri kümeleri hakkında varsayımlarda bulunmayan ve katı kısıtlar içermeyen yöntemlerin uygulama ve kullanım maliyetlerinin azalmış olmasıdır. SA, KEK, KA, DVM gibi, ana kütle üzerinde varsayım ve parametreleri karşılamaya gerek kalmaksızın mevcut veri kümeleri üzerinde analiz yapılmasını sağlayan yöntemler kredi değerlendirme çalışmalarının güncel akımını oluşturmaktadır. Öğrenen yöntemlere dair yazına 2000’li yılların başında sunulan RO da bu ikinci grup yöntemler arasında yer almaktadır ve kendisine biyoloji ve tıp gibi yoğun değişken kullanımı gerektiren alanlarda yapılan çalışmalarda sıklıkla başvurulmaktadır.

RO ağaç tabanlı yöntemler olarak adlandırılan grup içerisinde yer almaktadır. Tüm ağaç tabanlı yöntemlerin çalışma prensibi, tahmin değişken uzayının ayrıştırılması ile ortaya çıkan bölgelerde bulunan bağımlı değişken sınıf ya da değerlerine göre tahmin yapılmasına dayanmaktadır. En basit ağaç tabanlı yöntem olan KA, yüksek tahmin başarısı veren bölümlerle tahmin değişkenleri arasında hiyerarşik bir yapı kurar ve bu yapıyı basit şekilde görselleştirerek karar vericilere anlaşılır bilgi sağlar. Ancak fazlasıyla esnek bir yöntem olduğundan ve eğitilen veri kümesindeki gözlemlere aşırı hassasiyet gösterdiğinden yüksek varyans problemine maruz kalabilmektedir. Yüksek varyans sebebiyle de farklı eğitim kümeleri ile eğitildiğinde birbirinden uzak bağımlı değişken tahminleri ortaya çıkabilmektedir. Ağaç budaması ve torbalanmış ağaçlar gibi geliştirmeler ile söz konusu problemler giderilebilmektedir ancak baskın bir tahmin değişkeni ya da kendi arasında yüksek ilişki düzeyine sahip tahmin değişkenlerinin varlığı durumunda söz konusu geliştirmeler yeterli olamamaktadır. RO söz konusu olumsuzlukların hepsinin birden giderilmesi konusunda büyük başarılar göstermektedir. RO’ın öğrenme sürecinde çok sayıda ağaç oluşturulmakta ve her ağaç içerisindeki her bölümlerle karar, tüm değişkenler yerine yalnızca bir grup değişken ile verilmektedir. Böylece RO, birbiriyle düşük ilişki düzeyine sahip farklı ağaçlar ile bilgi elde edebilmekte ve tahminlerini buna göre yapabilmektedir. Kredi değerlendirmelerinde kullanılan ve başvurulara ait birbiriyle ilişkili olabilen çok sayıda değişken içeren veri kümelerinde RO büyük başarı potansiyeline sahiptir ancak ilgili yazındaki çalışmalar incelendiğinde RO yönteminin kullanım örneklerine nadiren rastlanılmaktadır.

Herhangi bir veri kümesinin doğru şekilde öğrenilmesinde ve sonrasında yapılacak tahminlerde yüksek başarı düzeylerine ulaşılmasında farklı yöntemlerin birlikte kullanımları büyük katkı sağlamaktadır. Özellikle değişken seçimine yönelik yöntemlerin entegrasyonu ile öğrenme süreçlerinin basitleştirilmesi ile yapılacak çıkarımların gücünün artması, benzer bilgi sağlayan değişkenlerin ortaya çıkaracakları çoklu bağlanım sorununun ortadan kaldırılması ve işlem süre/maliyetlerinin azaltılması sağlanabilir. Değişken seçimi teknikleri ya da içerisinde değişken seçimine yönelik mekanizmalar barındıran teknikler, özellikle az gözlem çok değişken içeren veri kümelerinin analizlerinde çözüm uzayı boyutunu azaltarak öğrenme performansını artırmaktadırlar. RO, öğrenme süreci içerisinde değişkenler arasında hiyerarşi oluşturarak değişken seçimini gömülü olarak gerçekleştirebilmektedir. Yine de bu çalışmada önerilen GA-TB birleştirilmiş algoritma kullanımı gibi ek sarmal değişken seçimi süreci dizaynları ile çok daha yüksek tahmin başarılarına erişilebilmektedir.

GA ve TB, meta-sezgiseller sınıfına giren ve optimizasyon çalışmalarında sıklıkla başvurulan iki popüler algoritmadır. Meta-sezgisel algoritmalar klasik optimizasyon yöntemlerinden farklı mekanizmalara sahiptirler ve özellikle karmaşık problemlerde anlamlı bir çözüme ulaşma konusunda oldukça başarılıdırlar. Bu algoritmalar, yapılan çalışmaya bağlı olarak tek ya da başka yöntemlerle birlikte/hibrit olarak kullanılarak tüm sürecin analiz performansını artırabilirler. Ancak meta-sezgisel algoritmalar, hangi biçimde kullanılırsa kullanılsın klasik matematiksel optimizasyon yöntemleri gibi her zaman en iyi çözümü garanti etmemektedirler. Tüm meta-sezgiseller, en iyi çözüm yerine en iyiye yakın çözümlerin kabulünü de içeren mekanizmalara sahiptir.

Meta-sezgisel algoritmaların ve özellikle GA'nın farklı alanlarda karşılaşılan birçok probleme yönelik oldukça başarılı uygulamaları bulunmaktadır. Algoritma mekanizmalarının genel yapısı ve farklı problemlere kolaylıkla uyarlanabilir olması meta-sezgisellerin en büyük kullanım avantajlarının başında gelmektedir. Çözüm uzayı tarama mekanizmalarının iyi anlaşılması ve problem gerekliliklerinin çözüm gösterimlerine doğru şekilde aktarılması, meta-sezgisel uygulamalarının başlıca katkısını oluşturmaktadır.

Bu çalışmada önerilen BA kullanımının birinci aşamasını oluşturan GA, biyolojik organizmaların zekâ, öğrenme ve adapte olma özelliklerinden ilham alınarak oluşturulmuştur ve organizmaların genleri ile taşıdığı faydalı bilgilerin sonraki nesillere aktarım mekanizmalarını taklit etmektedir. Buna göre mevcut bir popülasyon içerisinde güçlü genlere (çözümlere) sahip bireyler, doğal seçim yoluyla bu bilgiyi sonraki nesillere (yeni çözümlere) aktaracak ve bir süre sonra popülasyonda yalnızca güçlü olan bireyler (en iyi çözümler) kalacaktır. Bir bireyin güçlü olması onun, ele alınan problemin çözümünde başarılı olması anlamına gelmektedir. Optimizasyon problemlerinde başarı genellikle, belli bir hedef değer en küçük ya da en büyük olması ile ölçülür. Çok amaçlı karar verme problemlerinde olduğu gibi farklı çıktı hedefleri olabilmekle birlikte genel yaklaşım bu yöndedir. Kredi değerlendirme yazınında GA'nın son yıllarda değişken seçimi yapılması amacıyla başka yöntemler ile birlikte ya da hibrit olarak sıklıkla kullanıldığı ve tahmin başarılarına önemli katkılar yaptığı bilinmektedir.

BA'nın ikinci aşamasını oluşturan TB ise GA'dan farklı olarak çözüm uzayında bir popülasyon ile değil, tek bir çözüm üzerinden komşuluk oluşturarak tarama yapmaktadır. Benzetim, tavlama adı verilen ve katıların yüksek derecelere ısıtılması (başlangıç çözüm oluşturma), sonrasında kontrollü şekilde soğutulması sırasında daha tutarlı kristalize yapıya sahip hallerin ortaya çıkartılması (komşu çözümler) ve süreç sonunda en tutarlı yapıya ulaşılması (en iyi çözüm) işleminden esinlenilerek dizayn edilmiştir. TB'nin, güçlü bir meta-sezgisel algoritma olmasına ve başka alanlarda yoğun kullanım örnekleri bulunmasına rağmen kredi değerlendirme çalışmalarında örneklerine rastlanmamıştır. Bu açıdan kredi değerlendirmesi için yöntem olarak sunumu, mevcut yazına önemli bir katkı olarak görülmektedir.

GA ile TB arasında, çözüm uzayını tarama etkinlikleri bakımından farklar bulunmaktadır. GA'lar, prematüre yakınsama ya da yerel en iyiye takılma gibi problemlere maruz kalabilmektedir ve bu durumlar en iyi çözümün araştırılmasının önünde büyük engel oluşturmaktadır. Her ne kadar mutasyon bileşeni ile bu sorunların üstesinden gelmek mümkün olabilse de GA'lar, çözüm uzayının farklı alanlarına geçiş yapmakta TB kadar başarılı olamamaktadır. TB, mutasyon bileşeninden daha etkin bir atlatma mekanizması olan Metropolis seçim stratejisi sayesinde yerel en iyiye takılmış bir çözümü çözüm uzayında farklı noktalara taşıyabilir. Güçlü yoğunlaşma

mekanizmalarına sahip olan GA'yı güçlü çeşitlenme mekanizmasına sahip TB ile birlikte kullanarak her iki yöntemin de avantajlarından faydalanmak mümkündür. Kredi değerlendirme yazınında meta-sezgisel yöntemler ile birlikte kullanılan istatistiksel/makine öğrenme yöntemlerine dair örnekler olmakla birlikte farklı iki meta-sezgiselin dizayn avantajlarından faydalanılarak birleştirilmiş bir model oluşturulması ve bu modelin bir öğrenen yöntem ile hibritleştirilmesi örneği bulunmamaktadır.

BA-RO yöntemi, GA ve TB yöntemlerinin ayrı ayrı sahip oldukları zayıflıkları bertaraf etmeyi ve değişken seçimine yönelik kullanım avantajlarından faydalanarak RO yönteminin tahmin başarısını artırmayı hedeflemiştir. Yönteme göre öncelikle hibrit GA-RO modeli kullanılmış, en iyi tahmin performansını sağlayan çözüm TB'ne aktarılmış ve ardından hibrit TB-RO modeli kullanılmıştır. Öğrenen yöntemlerin eğitim süreçleri, eğitim kümesi büyüklüklerinden fazlasıyla etkilendiği için beş farklı test kümesi büyüklüğü belirlenmiştir. Gerek RO yönteminin gerekse meta-sezgisellerin rassal doğaları gereği anlamlı yorum yapılabilmesi için her büyüklük değerinde yirmi farklı eğitim kümesi oluşturulmuş ve tahmin performanslarının eğitim kümelerindeki gözlemlere göre ne düzeyde farklılaştığı araştırılmıştır. Belirli bir test kümesi büyüklüğündeki tahmin performansı, o büyüklükte oluşturulmuş tüm eğitim kümelerinin tahmin performanslarının ortalaması alınarak hesaplanmıştır. Önerilen BA'ın RO üzerindeki iyileştirme etkisini ölçmek için yukarıda bahsedilmiş eğitim ve test kümelerinde yalnızca RO yöntemi kullanılarak ayrı bir analiz gerçekleştirilmiş ve aralarında istatistiksel bir fark olup olmadığı araştırılmıştır. Analiz edilen veri kümesi, kredi değerlendirmesine yönelik yöntem önerisinde bulunan çalışmalarda sıklıkla kullanıldığından, BA-RO yöntem sonuçlarının diğer çalışmalardaki yöntemlerle karşılaştırılması imkânı da sağlanmıştır.

Yalnızca RO yöntemi ile gerçekleştirilen öğrenim performansının farklı test kümesi büyüklüklerinden etkilenmediği görülmüştür. Yöntem hassasiyetinin ölçüsü olan TİP-I kesinlik değerleri tüm test kümesi büyüklüklerinde düşük çıkmıştır. Söz konusu sorun, özgülük olarak da adlandırılan TİP-II kesinliklerin yüksek başarı değerleri karşısında daha da belirgin olmaktadır. TİP-I kesinlikteki bu düşük performansın, veri kümesinde bulunan "iyi" ve "kötü" gözlem sayılarındaki dengesizlikten kaynaklandığını söylemek mümkündür. "kötü" başvurulara ilişkin

yeterli sayıda ayırt edici örneğe sahip olmayan RO, bu sınıfa ait gözlemleri hatalı değerlendirme eğilimi göstermiştir.

Önerilen BA-RO yöntem bulguları, RO'ın hem genel tahmin hem de TİP-I kesinlik başarılarında anlamlı artış elde edildiğini göstermektedir. Tüm test kümesi büyüklükleri için tahmin başarıları ortalama %4.1 oranında artış göstermiştir. En fazla iyileşme, tüm gözlemlerin %10 ve %20'si ile oluşturulan test kümesi büyüklüklerinde ortaya çıkmıştır. En iyi tahmin başarısı, %10 oranında gözlem içeren test kümesi büyüklüğünde sağlanmıştır. Aynı test kümesi büyüklüğünde TİP-I ve TİP-II kesinlikler de en iyi değerlerine ulaşmıştır. Bu bulgular önerilen yöntemin eğitimi sırasında olabildiğince çok sayıda gözlem kullanılmasının daha iyi olduğunu göstermektedir. Çok sayıda gözlem ile eğitilen RO değişken seçimi uygulamasıyla birlikte, veri kümelerindeki ilişkileri aşırı uyum göstermeden ve genelleme yeteneğini kaybetmeden başarıyla öğrenmiştir. Ayrıca %10 ve %20'lik oranlarda test gözlemi içeren simülasyonlarda elde edilen tahmin hatalarının standart sapma değerlerinde yaşanan düşüşler de BA-RO genelleme performanslarının çok daha güvenilir olduğunu işaret etmektedir.

Temerrüde düşmüş kredi başvurularının tahmin başarılarında (TİP-I kesinlik) tüm test kümesi büyüklüklerinde artış sağlanmıştır. Kredi değerlendirmesinde temerrüde düşme riskine sahip başvuruların belirlenmesinin, genel tahmin başarısına ya da kaliteli kredilerin belirlenmesine göre çok daha önemli olduğu bilinmektedir. Buna göre önerilen yöntem problemlili başvuruların ortaya çıkartılması konusunda oldukça iyi bir katkı sağlamaktadır. Kimi test kümesi büyüklüklerinde, özgülük olarak adlandırılan TİP-II kesinliklerde yaşanan düşüşler ise, zaten yüksek düzeyde olan performansları içerisinde çok önemli gözükmemektedir. Analiz bulguları BA-RO kullanımının, RO yönteminin tek başına kullanımına göre tüm açılardan çok daha iyi sonuçlar verdiğini göstermektedir.

Aynı veri kümesini incelemiş yazındaki diğer çalışmalar ile karşılaştırıldığında da BA-RO yöntem performansı çok iyi bir noktada olduğu görülmüştür. BA-RO, tüm çalışmalar arasında yüksek performans elde edilen ilk dört çalışma arasında yer almaktadır. Söz konusu çalışmalarda kullanılan öğrenen yöntemlerin genellikle tek olarak ya da başka yöntemler ile birlikte veya GA gibi meta-sezgiseller ile hibrit olarak

kullanıldığı görülmektedir. İki meta-sezgiselin birlikte kullanıldığı BA-RO bu çalışmalar arasında farklı bir noktada bulunmaktadır.

Yöntem ile gerçekleştirilen değişken seçimi sonuçları, yöntemin eğitiminde tüm değişkenlerin kullanılması yerine daha az sayıda değişken içeren bir gruba başvurulmasının tahmin performansını artırdığını göstermektedir. Yine de bu durum, herhangi bir değişkenin diğerlerine göre çok daha fazla bilgi sağladığı ya da belirli değişkenlerin analiz dışında tutulması gerektiği gibi bir sonuca ulaştırmamaktadır. Farklı test ve eğitim kümesi büyüklüklerinde kimi değişkenlerin daha yüksek, kimilerinin de daha düşük oranlarda seçildiği görülmektedir. Genel seçim oranlarına bakıldığında müşterilere ait vadesiz mevduat hesap durumu, çekilen kredinin miktarı ve vadesi, müşteri kredi geçmişi gibi birkaç değişkenin biraz daha ön planda bulunmaktadır. Ancak bu ve benzeri çalışmalarda değişkenler arası etkileşimlerin ortaya çıkarttığı etkilerin tahmin performansında ne derece etkili olduğu kesin olarak bilinemediğinden, kullanılabilecek tüm değişkenlerin analizlere eklenmesi ve seçimin yönteme bırakılması daha doğrudur. Önerilen yöntemin, tahmin performansına etki etmeyen ya da düşüren değişkenleri hem BA ile adım adım hem de RO öğrenme süreci içerisinde ağaç oluşturma mekanizmalarıyla dışarıda tutma kabiliyeti yüksektir. Öğrenen yöntemler, öznel incelemeler ile tespit edilmesi neredeyse imkânsız olan tahmin değişkenleri ile bağımlı değişken arasındaki ilişkileri ortaya çıkartabilmekte ve bu bilgiyi kullanarak tahmin yapabilmektedirler. Bu açıdan önerilen yöntemin ve yazında kullanılan diğer yöntemlerin, insana bağlı inceleme ve karar süreçlerine göre üstünlüğü tartışılmazdır.

Alman tüketici veri kümesi, yazındaki çalışmalar ile performans karşılaştırması yapmak imkânı verse de önerilen yöntemin etkinliğinin sınırlarının denenmesi açısından yeterli olamamaktadır. Veri kümesinde, o günün mevcut ekonomik koşullarına ilişkin bilgi bulunmaması ve bu sebeple tüketicilerin kendilerinden kaynaklanmayan dışsal etkilerin incelemeye alınamaması yöntemin tüm tahmin kapasitesinin ortaya konmasını engellemektedir. Bu çerçevede ilerleyen çalışmalarda, bankalardan elde edilecek veri kümelerine kredi temerrüt olasılıklarını etkileyebileceği düşünülen tüm dışsal faktörlere ilişkin göstergelerin dahil edilmesi planlanmaktadır. Böylelikle kredi değerlendirmesinin temel hedefi olan temerrüt olasılıklarının hesaplanmasının yanında, sağlıklı kredi ortamlarının hangi koşullar

altında sağlanabileceğinin ortaya çıkartılmasına ilişkin de çalışmalar yürütmek mümkün olacaktır. Ne tür makro-ekonomik gelişmelerin kredi geri ödemelerini problemli hale getirdiği ortaya çıkartıldığında, sorunların önüne geçilmesi ya da etkilerinin azaltılması yönünde atılacak adımların da doğru şekilde belirlenmesi sağlanabilecektir. Örneğin aldığı krediyi değişken faiz oranı ile geri ödeyen ya da yabancı para cinsinden borçlanan taraflar, bu değişkenlerde yaşanan ani yükselişler karşısında büyük zararlara uğrayacak ve hatta geri ödeyememe sorunu ile karşı karşıya kalacaktır. Yaklaşan krizlerin önceden tespitine yönelik oluşturulmaya çalışılan erken uyarı sistemlerinin, kredi değerlendirme süreçlerine de entegrasyonu ile birçok olumsuzluğun önüne geçilebilir. Erken uyarı sistemlerinde amaç yaklaşan kriz ve değişimlerin yeterli bir süre öncesinde tespiti ve uygun aksiyomların alınmasıdır. Benzer yaklaşım, mevcut kredilerin takibinde ya da verilmek üzere olan kredilerin değerlendirilmesinde de kullanılabilir.

Bulaşma etkisi olarak adlandırılan olgunun yıkıcı etkileri, küresel finans krizi ve sonrasında Avro para bölgesinde yaşanan krizi ile birlikte net bir şekilde ortaya konmuştur. Karmaşık piyasalar ağı içerisinde çok sayıda yayılım kanalı bulunduğundan herhangi bir noktada ortaya çıkan olumsuzluğun kredi piyasaları üzerinde etkisi olmayacağını beklemek de doğru değildir. Kredi başvuruları ve mevcut kredilere ilişkin değişkenler zaman serisi analizlerinin de dahil olduğu, bu çalışmadaki gibi gelişmiş öğrenme yaklaşımları içerisinde sürekli olarak değerlendirilmelidir. Yalnızca makro-mikro ekonomik ve finans göstergelerinin değil, bu göstergelere ait belirsizliklerin de modellere dahil edilmesiyle büyük kazanımlar elde edilebilir. Tüm bu faktörlerin eklenmesiyle ortaya çıkacak olan ve çok sayıda değişken ve gözlem içeren veri kümeleri, önerilen BA-RO yöntemi ile yüksek başarı oranlarıyla analiz edilebilecektir.

Gerçekleştirilen ve ileride yapılması planlanan çalışmalar hem denetleyici ve düzenleyici kurumlar hem de son karar vericiler için fazlasıyla yararlı olacaktır. Otoriteler, değişen koşullar altında da sürekli bilgi sağlayan bir sisteme bağlı olarak çok daha anlamlı politikalar geliştirebilecektir. Böylelikle daha güvenilir ve sağlam bir finansal sistem dizaynına ulaşmak mümkün olabilir.

## KAYNAKÇA

Abellán, J. ve Castellano, J. G. (2017). A Comparative Study on Base Classifiers in Ensemble Methods for Credit Scoring. *Expert Systems with Applications*, 73: 1–10. doi:10.1016/j.eswa.2016.12.020

Acemoglu, D. ve Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty* (1. bs.). New York: Crown.

Addo, P. M., Guegan, D. ve Hassani, B. (2018). Credit Risk Analysis Using Machine and Deep Learning Models. *Risks*, 6(2): 38. doi:doi:10.3390/risks6020038

Aguilar-Rivera, R., Valenzuela-Rendón, M. ve Rodríguez-Ortiz, J. J. (2015). Genetic Algorithms and Darwinian Approaches in Financial Applications: A Survey. *Expert Systems with Applications*, 42(21): 7684–7697. doi:10.1016/j.eswa.2015.06.001

Ala'Raj, M. ve Abbod, M. F. (2016a). Classifiers Consensus System Approach for Credit Scoring. *Knowledge-Based Systems*, 104: 89–105. doi:10.1016/j.knosys.2016.04.013

Ala'Raj, M. ve Abbod, M. F. (2016b). A New Hybrid Ensemble Credit Scoring Model Based on Classifiers Consensus System Approach. *Expert Systems with Applications*, 64(C): 36–55. doi:10.1016/j.eswa.2016.07.017

Allen, B., Chan, K. K., Milne, A. ve Thomas, S. (2012). Basel III: Is the Cure Worse than the Disease? *International Review of Financial Analysis*, 25: 159–166.

Altınbaş, H. ve Biskin, O. T. (2015). Selecting Macroeconomic Influencers on Stock Markets by Using Feature Selection Algorithms. *Procedia Economics and Finance*, 30: 22–29. doi:10.1016/S2212-5671(15)01251-4



Altinbas, H., Pacelli, V. ve Sica, E. (2018). The Determinants of Sovereign Bond Yields in the EMU: New Empirical Evidence. *International Journal of Economics and Finance*, 10(5): 41.

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4): 589–609. doi:10.1111/j.1540-6261.1968.tb00843.x

Angelini, P., Clerc, L., Cúrdia, V., Gambacorta, L., Gerali, A., Locarno, A., ... Vlček, J. (2015). BASEL III: Long-term Impact on Economic Performance and Fluctuations. *The Manchester School*, 83(2): 217–251.

Antão, P. ve Lacerda, A. (2008). An Assessment of Capital Requirements under Basel II: The Portuguese Case. *Financial Stability Report/ 2008*, 191.

Atkinson, T., Luttrell, D. ve Rosenblum, H. (2013). How Bad Was It? The Costs and Consequences of the 2007–09 Financial Crisis. *Staff Papers*, (Temmuz).

Baesens, B., Setiono, R., Mues, C. ve Vanthienen, J. (2003). Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation. *Management Science*, 49(3): 312–329. doi:10.1287/mnsc.49.3.312.12739

Banking Act of 1933 (Glass-Steagall). (2013). *Federal Reserve History*. [https://www.federalreservehistory.org/essays/glass\\_steagall\\_act](https://www.federalreservehistory.org/essays/glass_steagall_act), (8.2.2018).

Basak, S., Kar, S., Saha, S., Khaidem, L. ve Dey, S. R. (2018). Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*. doi:10.1016/J.NAJEF.2018.06.013

Bellotti, T. ve Crook, J. (2009). Support Vector Machines for Credit Scoring and Discovery of Significant Features. *Expert Systems with Applications*, 36(2): 3302–

3308. doi:10.1016/j.eswa.2008.01.005

Benzin, A., Trück, S. ve Rachev, S. T. (2003). Approaches to Credit Risk in the New Basel Capital Accord. *Credit Risk* içinde (ss. 1–33). Springer.

BIS. (2000). *Principles for Management of Credit Risk*.

BIS. (2006). *Sound Credit Risk Assessment and Valuation for Loans*.

BIS. (2010a). *The Basel Committee's Response to the Financial Crisis: Report to the G20*.

BIS. (2010b). *The Basel Committee's Response to the Financial Crisis: Report to the G20*.

BIS. (2017a). Table F1.1. *Statistics Database*. <http://stats.bis.org/statx/srs/table/f1.1.>, (5.8.2017).

BIS. (2017b). History of the Basel Committee. [https://www.bis.org/bcbs/history.htm#basel\\_i](https://www.bis.org/bcbs/history.htm#basel_i), (6.12.2017).

Blanco, A., Pino-Mejías, R., Lara, J. ve Rayo, S. (2013). Credit Scoring Models for the Microfinance Industry Using Neural Networks: Evidence from Peru. *Expert Systems with Applications*, 40(1): 356–364. doi:10.1016/j.eswa.2012.07.051

Boulesteix, A. L., Janitza, S., Kruppa, J. ve König, I. R. (2012). Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. doi:10.1002/widm.1072

Breiman, L. (2001). Random Forests. *Machine learning*, 45(1): 5–32.

Brown, I. ve Mues, C. (2012). An Experimental Comparison of Classification Algorithms for Imbalanced Credit Scoring Data Sets. *Expert Systems with Applications*, 39(3): 3446–3453. doi:10.1016/j.eswa.2011.09.033

Cangürel, O. (2012). *Basel II Kapsamında Kredi Riskinin Ölçümünde Otorite Etkinliği: Türkiye için Alternatif Bir Öneri*. Türkiye Bankalar Birliği.

Cangürel, O., Güngör, S., Sevinç, V. U., Kayci, İ. ve Atalay, S. (2010). *Sorularla Basel III. Bankacılık Düzenleme Denetleme Kurumu*. Ankara.

Caprio, G. ve Klingebiel, D. (2003). *Episodes of Systemic and Borderline Banking Crises*. World Bank.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L. F., Nobrega, J. P. ve Oliveira, A. L. I. (2016). Computational Intelligence and Financial Markets: A Survey and Future Directions. *Expert Systems With Applications*, 55: 194–211. doi:10.1007/s00115-014-4101-z

Černý, V. (1985). Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm. *Journal of Optimization Theory and Applications*, 45(1): 41–51. doi:10.1007/BF00940812

Chang, C.-L., Jiménez-Martin, J.-Á., McAleer, M. ve Pérez-Amaral, T. (2011). Risk Management of Risk under the Basel Accord: Forecasting Value-at-Risk of VIX Futures. *Managerial Finance*, 37(11): 1088–1106.

Chaparala, A. ve Sajja, R. (2018). Optimization Using Evolutionary Metaheuristic Techniques: A Brief Review. *Brazilian Journal of Operations & Production Management*, 15(1): 44–53.

Chen, F. L. ve Li, F. C. (2010). Combination of Feature Selection Approaches with

SVM in Credit Scoring. *Expert Systems with Applications*, 37(7): 4902–4909. doi:10.1016/j.eswa.2009.12.025

Chen, N., Ribeiro, B. ve Chen, A. (2016). Financial Credit Risk Assessment: a Recent Review. *Artificial Intelligence Review*, 45(1): 1–23. doi:10.1007/s10462-015-9434-x

Claessens, S. ve Kodres, M. L. E. (2014). *The Regulatory Responses to the Global Financial Crisis: Some Uncomfortable Questions*. International Monetary Fund.

Crotty, J. (2009). Structural Causes of the Global Financial Crisis: A Critical Assessment of the “New Financial Architecture”. *Cambridge Journal of Economics*, 33(4): 563–580. doi:10.1093/cje/bep023

Cubiles-De-La-Vega, M.-D., Blanco-Oliver, A., Pino-Mejías, R. ve Lara-Rubio, J. (2013). Improving the Management of Microfinance Institutions by Using Credit Scoring Models Based on Statistical Learning Techniques. *Expert Systems with Applications*, 40(17): 6910–6917. doi:10.1016/j.eswa.2013.06.031

Dall, N. (1992). *A review of the credit application and assessment process from a consumer perspective*. Edith Cowan University.

Danenas, P. ve Garsva, G. (2015). Selection of Support Vector Machines Based Classifiers for Credit Risk Domain. *Expert Systems with Applications*, 42(6): 3194–3204. doi:10.1016/j.eswa.2014.12.001

De Luca, G., Riviuccio, G. ve Zuccolotto, P. (2010). Combining Random Forest and Copula Functions: A Heuristic Approach for Selecting Assets From a Financial Crisis Perspective. *Intelligent Systems in Accounting, Finance & Management*, 17(2): 91–109. doi:10.1002/isaf.315

Decamps, J.-P., Rochet, J.-C. ve Roger, B. (2004). The Three Pillars of Basel II:

Optimizing the Mix. *Journal of Financial Intermediation*, 13(2): 132–155.

Desai, V. S., Conway, D. G., Crook, J. N. ve Overstreet Jr., G. A. (1997). Credit-scoring Models in the Credit-union Environment Using Neural Networks and Genetic Algorithms. *IMA Journal of Mathematics Applied in Business and Industry*, 8: 323–346.

Desai, V. S., Crook, J. N. ve Overstreet, G. A. (1996). A Comparison of Neural Networks and Linear Scoring Models in the Credit Union Environment. *European Journal of Operational Research*, 95(1): 24–37. doi:10.1016/0377-2217(95)00246-4

Dinh, T. H. T. ve Kleimeier, S. (2007). A Credit Scoring Model for Vietnam's Retail Banking Market. *International Review of Financial Analysis*, 16(5): 471–495. doi:10.1016/j.irfa.2007.06.001

Dionne, G. (2013). Risk Management: History, Definition, and Critique. *Risk Management and Insurance Review*, 16(2): 147–166.

Elliott, D. J. (2010). *Basel III, the Banks, and the Economy*. Brookings Institutions.

Ferguson, R. W. (2003). Capital Standards for Banks: The Evolving Basel Accord. *Federal Reserve Bulletin*, 89(9): 395–405.

Fidrmuc, J. ve Lind, R. (2018). Macroeconomic Impact of Basel III: Evidence from a Meta-Analysis. *Journal of Banking & Finance*.

Flood, M. H. (2017). *Early Identification of High-Risk Credit Card Customers Based on Behavioral Data*. Norwegian University of Science and Technology.

Florez-Lopez, R. ve Ramon-Jeronimo, J. M. (2015). Enhancing Accuracy and Interpretability of Ensemble Strategies in Credit Risk Assessment. A Correlated-

adjusted Decision Forest Proposal. *Expert Systems with Applications*, 42(13): 5737–5753. doi:10.1016/j.eswa.2015.02.042

Fortmann-Roe, S. (2012). Understanding the Bias-variance Tradeoff. <http://scott.fortmann-roe.com/docs/BiasVariance.html>, (12.9.2018).

Fund, I. M. (2013a). *Key Aspects of Macroprudential Policy, Board Paper*.

Fund, I. M. (2013b). *Key Aspects of Macroprudential Policy, Background Paper*.

Gendreau, M. ve Potvin, J. Y. (2010). *Handbook of Metaheuristics* (2. bs.). New York: Springer. doi:10.1007/978-1-4614-1900-6

Georg, C.-P. (2011). *Basel III and Systemic Risk Regulation- What Way Forward?*

Girginer, N. (2008). Ticari Kredi Taleplerinin Değerlendirmesine Çok Kriterli Yaklaşım: Özel ve Devlet Bankası Karşılaştırması. *Muhasebe ve Finansman Dergisi*, (37): 132–142.

Gümüşçü, A., Aydılek, İ. B. ve Taşaltın, R. (2016). Mikro-dizilim Veri Sınıflandırmasında Öznitelik Seçme Algoritmalarının Karşılaştırılması. *Harran Üniversitesi Mühendislik Dergisi*, 1: 1–7.

Gürel, E., Gürel, E. B. B. ve Demir, N. (2012). Basel III Kriterleri. *Bankacılık ve Sigortacılık Araştırmalar Dergisi*, 1(3–4).

Harari, Y. N. (2014). *Sapiens: A Brief History of Humankind*. London: Harvill Secker.

Harrell, F. E. (2001). *Regression Modeling Strategies, with Applications to Linear Models, Survival Analysis and Logistic Regression*. Springer.

Harris, T. (2015). Credit Scoring Using the Clustered Support Vector Machine. *Expert Systems with Applications*, 42(2): 741–750. doi:10.1016/j.eswa.2014.08.029

Hastie, T., Tibshirani, R. ve Friedman, J. (2009). The Elements of Statistical Learning. *Bayesian Forecasting and Dynamic Models*, 1: 1–694. doi:10.1007/b94608

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation* (2. bs.). NJ: Prentice Hall.

Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT press.

Hsieh, N. C. (2005). Hybrid Mining Approach in the Design of Credit Scoring Models. *Expert Systems with Applications*, 28(4): 655–665. doi:10.1016/j.eswa.2004.12.022

Huang, C.-L., Chen, M.-C. ve Wang, C.-J. (2007). Credit Scoring with a Data Mining Approach Based on Support Vector Machines. *Expert Systems with Applications*, 33(4): 847–856. doi:10.1016/j.eswa.2006.07.007

Huang, Z., Chen, H., Hsu, C. J., Chen, W. H. ve Wu, S. (2004). Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study. *Decision Support Systems*, 37(4): 543–558. doi:10.1016/S0167-9236(03)00086-1

İç, Y. T. ve Yurdakul, M. (2000). Analitik Hiyerarisi Süreci (AHS) Yöntemini Kullanan Bir Kredi Değerlendirme Sistemi. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 15(1): 1–14.

Jadhav, S., He, H. ve Jenkins, K. (2018). Information Gain Directed Genetic Algorithm Wrapper Feature Selection for Credit Rating. *Applied Soft Computing Journal*, 69: 541–553. doi:10.1016/j.asoc.2018.04.033

James, G., Witten, D., Hastie, T. ve Tibshirani, R. (2013). *An introduction to Statistical Learning* (C. 112). New York: springer. doi:10.1007/978-1-4614-7138-7

James, G., Witten, D., Hastie, T. ve Tibshirani, R. (2017). ISLR: Data for an Introduction to Statistical Learning with Applications in R. <https://cran.r-project.org/package=ISLR>, (12.2.2017).

Jin, C., Jin, S. W. ve Qin, L. N. (2012). Attribute Selection Method Based on a Hybrid BPNN and PSO Algorithms. *Applied Soft Computing Journal*, 12(8): 2147–2155. doi:10.1016/j.asoc.2012.03.015

Kar, A. K. (2016). Bio Inspired Computing - A Review of Algorithms and Scope of Applications. *Expert Systems with Applications*, 59: 20–32. doi:10.1016/j.eswa.2016.04.018

Karels, G. V ve Prakash, A. J. (1987). Multivariate Normality and Forecasting of Business Bankruptcy. *Journal of Business Finance & Accounting*, 14(4): 573–593.

Kaynar, O., Arslan, H., Görmez, Y. ve Işık, Y. E. (2018). Makine Öğrenmesi ve Öznitelik Seçim Yöntemleriyle Saldırı Tespiti. *Bilişim Teknolojileri Dergisi*, 11(2): 175–185. doi:10.17671/gazibtd.368583

Kennedy, J. (2011). Particle Swarm Optimization. *Encyclopedia of machine learning* içinde (ss. 760–766). Springer US.

Keskintürk, T. (2007). Portföy Seçiminde Markowitz Modeli İçin Yeni Bir Genetik Algoritma Yaklaşımı. *Yönetim*, 18(56): 78–90.

Khashman, A. (2010). Neural Networks for Credit Risk Evaluation: Investigation of Different Neural Models and Learning Schemes. *Expert Systems with Applications*, 37(9): 6233–6239. doi:10.1016/j.eswa.2010.02.101



Kirkpatrick, S., Gelatt, C. D. ve Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, 220(4598): 671–680. doi:10.1126/science.220.4598.671

Koutanaei, F. N., Sajedi, H. ve Khanbabaei, M. (2015). A Hybrid Data Mining Model of Feature Selection Algorithms and Ensemble Learning Classifiers for Credit Scoring. *Journal of Retailing and Consumer Services*, 27: 11–23. doi:10.1016/j.jretconser.2015.07.003

Kozeny, V. (2015). Genetic Algorithms for Credit Scoring: Alternative Fitness Function Performance Comparison. *Expert Systems with Applications*, 42: 2998–3004.

Kruppa, J., Schwarz, A., Arminger, G. ve Ziegler, A. (2013). Consumer Credit Risk: Individual Probability Estimates Using Machine Learning. *Expert Systems with Applications*, 40(13): 5125–5131. doi:10.1016/j.eswa.2013.03.019

Kruppa, J., Ziegler, A. ve König, I. R. (2012). Risk Estimation and Risk Prediction Using Machine-Learning Methods. *Human Genetics*, 131(10): 1639–1654. doi:10.1007/s00439-012-1194-y

Kuhn, M. ve Johnson, K. (2013). *Applied Predictive Modeling* (26. bs.). New York: Springer.

Kumar, M. ve Thenmozhi, M. (2014). Forecasting Stock Index Returns Using ARIMA-SVM, ARIMA-ANN, and ARIMA-Random Forest Hybrid Models. *International Journal of Banking, Accounting and Finance*, 5(3): 284–308.

Lall, R. (2009). Why Basel II Failed and Why any Basel III is Doomed. *Global Economic Governance Programme, GEC Working Paper*, 52.

Lee, T.-S. ve Chen, I.-F. (2005). A Two-stage Hybrid Credit Scoring Model Using Artificial Neural Networks and Multivariate Adaptive Regression Splines. *Expert*

*Systems with Applications*, 28(4): 743–752. doi:10.1016/j.eswa.2004.12.031

Lee, T. S., Chiu, C. C., Lu, C. J. ve Chen, I. F. (2002). Credit Scoring Using the Hybrid Neural Discriminant Technique. *Expert Systems with Applications*, 23(3): 245–254. doi:10.1016/S0957-4174(02)00044-1

Lessmann, S., Baesens, B., Seow, H. V. ve Thomas, L. C. (2015). Benchmarking State-of-the-art Classification Algorithms for Credit scoring: An Update of Research. *European Journal of Operational Research*, 247(1): 124–136. doi:10.1016/j.ejor.2015.05.030

Liao, S. H., Chu, P. H. ve Hsiao, P. Y. (2012). Data Mining Techniques and Applications - A Decade Review from 2000 to 2011. *Expert Systems with Applications*, 39(12): 11303–11311. doi:10.1016/j.eswa.2012.02.063

Liaw, A. ve Wiener, M. (2002). Classification and Regression by RandomForest. *R News*, 2(3): 18–22.

Lichman, M. (2013). UCI Machine Learning Repository. Irvine, CA. <http://archive.ics.uci.edu/ml>, (10.12.2015).

Liu, C., Chan, Y., Alam Kazmi, S. H. ve Fu, H. (2015). Financial Fraud Detection Model: Based on Random Forest. *International Journal of Economics and Finance*, 7(7): 178–188. doi:10.5539/ijef.v7n7p178

Louzada, F., Ara, A. ve Fernandes, G. B. (2016). Classification Methods Applied to Credit Scoring: Systematic Review and Overall Comparison. *Surveys in Operations Research and Management Science*, 21(2): 117–134. doi:10.1016/j.sorms.2016.10.001

Lu, H., Liyan, H. ve Hongwei, Z. (2013). Credit Scoring Model Hybridizing Artificial

Intelligence with Logistic Regression. *Journal of Networks*, 8(1): 253–261.  
doi:10.4304/jnw.8.1.253-261

Maddison, A. (2001). *The World Economy: A Millennial Perspective*. *Oecd*.  
doi:10.1787/9789264189980-en

Maheshwari, A. ve Smid, M. (2017). *Introduction to Theory of Computation*. Ottawa:  
School of Computer Science, Carleton University.

Marcano-Cedeño, A., Marin-de-la-Barcelona, J., Jimenez-Trillo, J., Piñuela, J. ve  
Andina, D. (2011). Artificial Metaplasticity Neural Network Applied to Credit  
Scoring. *International Journal of Neural Systems*, 21(04): 311–317.  
doi:10.1142/S0129065711002857

Maringer, D. G. (2005). *Portfolio Management with Heuristic Optimization*. Springer  
US.

Mermod, A. Y. ve Ceran, M. (2011). Basel III Doğrultusunda Bankacılık Riskleri ve  
Sermaye Yeterliliği; Türk Bankacılık Sektörü Üzerine Karşılaştırmalı Bir Analiz.  
*Finansal Araştırmalar ve Çalışmalar Dergisi*, 2(4): 29–38.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. ve Teller, E.  
(1953). Equation of State Calculations by Fast Computing Machines. *The journal of  
chemical physics*, 21(6): 1087–1092.

Obadan, M. I. (2006). Globalization of Finance and the Challenge of National  
Financial Sector Development. *Journal of Asian Economics*, 17(2): 316–332.  
doi:10.1016/j.asieco.2005.11.002

Ojo, M. (2010). Basel III and Responding to the Recent Financial Crisis: Progress  
Made by the Basel Committee in Relation to the Need for Increased Bank Capital and

Increased Quality of Loss Absorbing Capital.

Oreski, S., Oreski, D. ve Oreski, G. (2012). Hybrid System with Genetic Algorithm and Artificial Neural Networks and Its Application to Retail Credit Risk Assessment. *Expert Systems with Applications*, 39(16): 12605–12617. doi:10.1016/j.eswa.2012.05.023

Oreski, S. ve Oreski, G. (2014). Genetic Algorithm-based Heuristic for Feature Selection in Credit Risk Assessment. *Expert Systems with Applications*, 41(4): 2052–2064. doi:10.1016/j.eswa.2013.09.004

Özdemir, M. (2011). Genetik Algoritma Kullanılarak Portföy Seçimi. *Iktisat Isletme ve Finans*, 26(299): 43–66. <https://ideas.repec.org/a/iif/iifjrn/v26y2011i299p43-66.html> adresinden erişildi.

Patrick, H. T. (1966). Financial Development and Economic Growth in Underdeveloped Countries. *Economic development and Cultural change*, 14(2): 174–189.

Penikas, H. (2015). History of Banking Regulation as Developed by the Basel Committee on Banking Supervision in 1974-2014 (Brief Overview). *Estabilidad Financiera*, 28: 9–48.

R, C. T. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. doi:10.1007/978-3-540-74686-7

Reichert, A. K., Cho, C. C. ve Wagner, G. M. (1983). An Examination of the Conceptual Issues Involved in Developing Credit-scoring Models. *Journal of Business and Economic Statistics*, 1(2): 101–141. doi:10.1080/07350015.1983.10509329

Rossignolo, A. F., Fethi, M. D. ve Shaban, M. (2013). Market Crises and Basel Capital

Requirements: Could Basel III have been Different? Evidence from Portugal, Ireland, Greece and Spain (PIGS). *Journal of Banking & Finance*, 37(5): 1323–1339. doi:<https://doi.org/10.1016/j.jbankfin.2012.08.021>

Salappa, A., Doumpos, M. ve Zopounidis, C. (2007). Feature Selection Algorithms in Classification Problems: An Experimental Evaluation. *Optimisation Methods and Software*, 22(1): 199–212.

Sarica, A., Cerasa, A. ve Quattrone, A. (2017). Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer’s Disease: A Systematic Review. *Frontiers in Aging Neuroscience*, 9(329): 1–12. doi:[10.3389/fnagi.2017.00329](https://doi.org/10.3389/fnagi.2017.00329)

Slovik, P. ve Cournède, B. (2011). *Macroeconomic Impact of Basel III* ( No: 844).

Soler-Dominguez, A., Juan, A. A. ve Kizys, R. (2017). A Survey on Financial Applications of Metaheuristics. *ACM Computing Surveys*, 50(1): 1–23. doi:[10.1145/3054133](https://doi.org/10.1145/3054133)

Sorensen, K., Sevaux, M. ve Glover, F. (2018). A History of Metaheuristics. *Handbook of Heuristics* içinde (ss. 791–808). Springer. doi:[10.1007/978-3-319-07153-4\\_4-1](https://doi.org/10.1007/978-3-319-07153-4_4-1)

Šušteršič, M., Mramor, D. ve Zupan, J. (2009). Consumer Credit Scoring Models with Limited Data. *Expert Systems with Applications*, 36(3): 4736–4744. doi:[10.1016/j.eswa.2008.06.016](https://doi.org/10.1016/j.eswa.2008.06.016)

Sylvester, E. V. A., Bentzen, P., Bradbury, I. R., Clément, M., Pearce, J., Horne, J. ve Beiko, R. G. (2018). Applications of Random Forest Feature Selection for Fine-scale Genetic Population Assignment. *Evolutionary Applications*, 11(2): 153–165. doi:[10.1111/eva.12524](https://doi.org/10.1111/eva.12524)

Taylor, B. J. ve Darrah, M. A. (2005). Rule Extraction as a Formal Method for the Verification and Validation of Neural Networks. *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, 5: 2915–2920.

Team, Rs. (2015). RStudio: Integrated Development Environment for R. Boston, MA: RStudio, Inc.

Valletta, J. J., Torney, C., Kings, M., Thornton, A. ve Madden, J. (2017). Applications of Machine Learning in Animal Behaviour Studies. *Animal Behaviour*, 124: 203–220. doi:10.1016/j.anbehav.2016.12.005

Van Sang, H., Nam, N. H. ve Nhan, N. D. (2016). A Novel Credit Scoring Prediction Model Based on Feature Selection Approach and Parallel Random Forest. *Indian Journal of Science and Technology*, 9(20): 1–6. doi:10.17485/ijst/2016/v9i20/92299

Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F. ve Fouilloy, A. (2017). Machine Learning Methods for Solar Radiation Forecasting: A Review. *Renewable Energy*, 105: 569–582. doi:10.1016/j.renene.2016.12.095

Wang, G., Ma, J., Huang, L. ve Xu, K. (2012). Two Credit Scoring Models Based on Dual Strategy Ensemble Trees. *Knowledge-Based Systems*, 26: 61–68. doi:10.1016/j.knosys.2011.06.020

WB. (2017a). GDP data. *Statistics Database*. <http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?end=2016&start=2006&view=chart>, (5.8.2017).

WB. (2017b). Unemployment data. *Statistics Database*. <http://data.worldbank.org/indicator/SL.UEM.TOTL.ZS?end=2016&start=2006>, (5.8.2017).

WB. (2017c). GDP Growth (annual %). *Statistics Database*.  
<https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG>, (6.8.2017).

West, D. (2000). Neural Network Credit Scoring Models. *Computers and Operations Research*, 27(11–12): 1131–1152. doi:10.1016/S0305-0548(99)00149-5

Yakut, E. ve Çankal, A. (2016). Çok Amaçlı Genetik Algoritma ve Hedef Programlama Metotlarını Kullanarak Hisse Senedi Portföy Optimizasyonu: BİST-30’da Bir Uygulama. *Business and Economics Research Journal*, 7(2): 43.

Yu, L., Yao, X., Wang, S. ve Lai, K. K. (2011). Credit Risk Evaluation Using a Weighted Least Squares SVM Classifier with Design of Experiment for Parameter Selection. *Expert Systems with Applications*, 38(12): 15392–15399. doi:10.1016/j.eswa.2011.06.023

Yücel, F. ve Kalyoncu, H. (2011). Finansal Krizlerin Öncü Göstergeleri ve Ülke Ekonomilerini Etkileme Kanalları: Türkiye Örneği. *Maliye Dergisi*, 159: 53–69. [http://dergiler.sgb.gov.tr/calismalar/maliye\\_dergisi/yayinlar/md/159/FatihYÜCEL.pdf](http://dergiler.sgb.gov.tr/calismalar/maliye_dergisi/yayinlar/md/159/FatihYÜCEL.pdf) adresinden erişildi.

Zäpfel, G. ve Braune, R. (2010). *Metaheuristic Search Concepts: A Tutorial with Applications to Production and Logistics*. Springer Science & Business Media.

Zeren, F. ve Baygın, M. (2015). Genetik Algoritmalar ile Optimal Portföy Seçimi: BİST-30 Örneği. *Journal of Business Research Turk*, 7(1): 309–324.

Zhang, D. Z. D., Hifi, M., Chen, Q. C. Q. ve Ye, W. Y. W. (2008). A Hybrid Credit Scoring Model Based on Genetic Programming and Support Vector Machines. *2008 Fourth International Conference on Natural Computation*, 7: 8–12. doi:10.1109/ICNC.2008.205

Zhang, D. Z. D., Leung, S. C. H. ve Ye, Z. Y. Z. (2008). A Decision Tree Scoring Model Based on Genetic Algorithm and K-Means Algorithm. *2008 Third International Conference on Convergence and Hybrid Information Technology*, 1: 1043–1047. doi:10.1109/ICCIT.2008.110

Zhao, Z., Xu, S., Kang, B. H., Kabir, M. M. J., Liu, Y. ve Wasinger, R. (2015). Investigation and Improvement of Multi-layer Perception Neural Networks for Credit Scoring. *Expert Systems with Applications*, 42(7): 3508–3516. doi:10.1016/j.eswa.2014.12.006

Zheng, C., Rahman, M. M., Begum, M. ve Ashraf, B. N. (2017). Capital Regulation, the Cost of Financial Intermediation and Bank Profitability: Evidence from Bangladesh. *Journal of Risk and Financial Management*, 10(2): 9.

Zopounidis, C., Doumpos, M. ve Niklis, D. (2018). Financial Decision Support: An Overview of Developments and Recent Trends. *EURO Journal on Decision Processes*, 6(1–2): 63–76. doi:10.1007/s40070-018-0078-3





# **EKLER**

## **EK 1: Yirminci Yüzyılda Yaşanmış Ekonomik ve Finansal Krizler Listesi**

- 1921 – Emtia Fiyatları Çöküşü.
- 1923 – Almanya’da aşırı-enflasyon
- 1929 – Büyük Depresyon
- 1931 – Birleşik Krallık, Japonya, Almanya ve Avusturya’da finansal kriz
- 1966 – ABD kredi krizi.
- 1973 – Dünya petrol krizi
- 1982 - Küresel kredi daralması
- 1987 – Tahvil ve senet piyasası çöküşü
- 1989 – Sorunlu tahvil krizi.
- 1995 – Meksika finans krizi
- 1997 – Asya finans krizi
- 1998 – Rusya’nın temerrüde düşmesi
- 2001 - Sorunlu tahvil krizi.
- 2001 – Arjantin’de kriz
- 2001 – Türkiye’de kriz
- 2002 – Brezilya’da tahvil piyasası krizi
- 2007 – ABD emlak krizi
- 2009- Avro bölgesi hükümet borç krizi

## EK 2: Alman Tüketici Kredi Veri Kümesi, Gözlemlere Ait Değişkenler

### Değişken 1: (Nitel)

Vadesiz mevduat hesap durumu:

A11 : ... < 0 Alman Markı

A12 :  $0 \leq \dots < 200$  Alman Markı

A13 : ...  $\geq 200$  Alman Markı

A14 : Hesabı yok

### Değişken 2: (Sayısal)

Kredi vadesi- aylık

### Değişken 3: (Nitel)

Kredi Geçmişi:

A30 : Kredi kullanmamış/ tüm krediler zamanında ödenmiş

A31 : Bu bankadaki tüm krediler zamanında ödenmiş

A32 : Mevcut krediler şimdiye kadar zamanında ödenmiş

A33 : Geçmiş ödemelerde gecikmeler yaşanmış

A34 : Kritik hesap/ bu banka hariç diğerlerinde kredi hesabı var

### Değişken 4: (Nitel)

Amaç

A40 : Araba (yeni)

A41 : Araba (kullanılmış)

A42 : Mobilya/ekipman

A43 : Radyo/televizyon

A44 : Yerli cihazlar

A45 : Tamirat

A46 : Eğitim

A47 : Tatil

A48 : Yeniden eğitim

A49 : Ticari

A410 : Diğer

Değişken 5: (Sayısal)

Kredi miktarı

Değişken 6: (Nitel)

Tasarruf hesabı/tahviller

A61 : ... < 100 Alman Markı

A62 : 100 <= ... < 500 Alman Markı

A63 : 500 <= ... < 1000 Alman Markı

A64 : .. >= 1000 Alman Markı

A65 : Bilinmiyor/tasarruf hesabı yok

Değişken 7: (Nitel)

Şu anki çalışma süresi

A71 : işsiz

A72 : ... < 1 yıl

A73 : 1 <= ... < 4 yıl

A74 : 4 <= ... < 7 yıl

A75 : ... >= 7 yıl

Değişken 8: (Sayısal)

Harcanabilir gelirin yüzdesi olarak taksit oranı

Değişken 9: (Nitel)

Medeni durum ve cinsiyet

A91 : Erkek : Boşanmış/ayrılmış

A92 : Kadın : Boşanmış /ayrılmış /evli

A93 : Erkek : Bekar

A94 : Erkek : Evli/dul

A95 : Kadın : Bekar

Değişken 10: (Nitel)

Ortak başvuru/ kefiller

A101 : Yok

A102 : Ortak başvuran var

A103 : Kefili var

Değişken 11: (Sayısal)

Mevcut ikametgâh süresi

Değişken 12: (Nitel)

Sahip olunan mülkler

A121 : Gayrimenkul

A122 : Eğer A121 değilse : Mortgage/hayat sigortası sahibi

A123 : Eğer A121/A122 değilse : araba veya diğer, Değişken 6'da yoksa

A124 : Bilinmiyor / mülkü yok

Değişken 13: (Sayısal)

Yaş

Değişken 14: (Nitel)

Diğer ödeme planları

A141 : Banka

A142 : Mağazalar

A143 : Mevcut değil

Değişken 15: (Nitel)

Barınma

A151 : Kira

A152 : Evin sahibi

A153 : Ücretsiz

Değişken 16: (Sayısal)

Bankadaki mevcut kredi sayısı

Değişken 17: (Nitel)

İş durumu

A171 : İşsiz/kalifiye değil- yerli değil

A172 : Kalifiye değil - yerli

A173 : Kalifiye çalışan/memur

A174 : Yönetici/iş yeri sahibi/yüksek düzeyde kalifiye çalışan/memur

Değişken 18: (Sayısal)

Bakmakla yükümlü olduğu kişi sayısı

Değişken 19: (Nitel)

Telefon

A191 : Mevcut değil

A192 : Mevcut, müşteri adına kayıtlı

Değişken 20: (Nitel)

Yabancı çalışan

A201 : Evet

A202 : Hayır

### EK 3: Değişkenlerin Betimleyici İstatistikleri

Değişken 1: (Nitel)

Vadesiz mevduat hesap durumu gözlem sayıları:

A11:274
A12:269
A13: 63
A14:394

Değişken 2: (Sayısal)

Kredi vadesi- aylık

Min.: 4.0
1. Çeyrek.: 12.0
Medyan: 18.0
Ortalama: 20.9
3. çeyrek.:24.0
Maks.: 72.0

Değişken 3: (Nitel)

Kredi Geçmiş gözlem sayısı:

A30: 40
A31: 49
A32:530
A33: 88
A34:293

Değişken 4: (Nitel)

Amaç gözlem sayısı:

A43 :280
A40 :234
A42 :181
A41 :103
A49 : 97
A46 : 50
(diğer): 55

Değişken 5: (Sayısal)

Kredi miktarı

Min.: 250
1. Çeyrek: 1366
Medyan: 2320
Ortalama: 3271
3. Çeyrek: 3972
Maks.:18424

Değişken 6: (Nitel)

Tasarruf hesabı/tahviller gözlem sayısı:

A61:603
A62:103
A63: 63
A64: 48
A65:183

Değişken 7: (Nitel)

Şu anki çalışma süresi gözlem sayısı:

A71: 62
A72:172
A73:339
A74:174
A75:253

Değişken 8: (Sayısal)

Harcanabilir gelirin yüzdesi olarak taksit oranı

Min.: 1
1. Çeyrek: 2
Medyan: 3
Ortalama: 2.973
3. Çeyrek:4
Maks.:4



Değişken 9: (Nitel)

Medeni durum ve cinsiyet gözlem sayısı:

A91: 50
A92:310
A93:548
A94: 92

Değişken 10: (Nitel)

Ortak başvuru/ kefiller gözlem sayısı:

A101:907
A102: 41
A103: 52

Değişken 11: (Sayısal)

Mevcut ikametgâh süresi

Min. :1
1. Çeyrek: 2
Medyan: 3
Ortalama: 2.845
3. Çeyrek: 4
Maks.: 4

Değişken 12: (Nitel)

Sahip olunan mülkler gözlem sayısı:

A121:282
A122:232
A123:332
A124:154

Değişken 13: (Sayısal)

Yaş

Min. :19.00
1. Çeyrek:27.00
Medyan:33.00
Ortalama: 35.55
3. Çeyrek:42.00
Maks.:75.00

Değişken 14: (Nitel)

Diğer ödeme planları gözlem sayısı:

A141:139
A142: 47
A143:814

Değişken 15: (Nitel)

Barınma gözlem sayısı:

A151:179
A152:713
A153:108

Değişken 16: (Sayısal)

Bankadaki mevcut kredi sayısı

Min. :1
1. Çeyrek:1
Medyan :1
Ortalama: 1.407
3. Çeyrek: 2
Maks.: 4

Değişken 17: (Nitel)

İş durumu gözlem sayısı:

A171: 22
A172:200
A173:630
A174:148

Değişken 18: (Sayısal)

Bakmakla yükümlü olduğu kişi sayısı

Min. :1
1. Çeyrek:1
Medyan :1
Ortalama:1.155
3. Çeyrek:1
Maks.:2

Değişken 19: (Nitel)

Telefon gözlem sayısı:

A191:596
A192:404

Değişken 20: (Nitel)

Yabancı çalışan gözlem sayısı:

A201:963
A202: 37