



**T.C.
HALIÇ ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**VARLIK YÖNETİMİ ŞİRKETİNDE SINIFLANDIRMA
YÖNTEMİYLE BORÇ GERİ ÖDEME EĞİLİMİ ANALİZİ**

YÜKSEK LİSANS TEZİ

Hazırlayan

Can TAŞCI

Danışman

Prof. Dr. Mübariz EMİNLİ

İstanbul – 2019

**T.C.
HALIÇ ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**VARLIK YÖNETİMİ ŞİRKETİNDE SINIFLANDIRMA
YÖNTEMİYLE BORÇ GERİ ÖDEME EĞİLİMİ ANALİZİ**

YÜKSEK LİSANS TEZİ

**Hazırlayan
Can TAŞCI**

**Danışman
Prof. Dr. Mübariz EMİNLİ**

İstanbul – 2019

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Programı Öğrencisi Can TAŞCI tarafından hazırlanan “*Varlık Yönetimi Şirketinde Sınıflandırma Yöntemiyle Borç Geri Ödeme Eğilimi Analizi*” konulu çalışması jürimizce Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 05.07.2019

(Jüri Üyesinin Ünvanı, Adı, Soyadı ve Kurumu):

İmzası

Jüri Üyesi : Prof.Dr. Mübariz EMİNLİ
: Haliç Üniversitesi (Danışman)

Jüri Üyesi : Dr. Öğretim Üyesi Zeynep TURGUT
: Haliç Üniversitesi

Jüri Üyesi : Dr. Öğr. Üyesi Selçuk SEVGİN
: İstanbul-Cerrahpaşa Üniversitesi

Bu tez Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından uygun görülmüş ve Enstitü Yönetim Kurulunun kararıyla kabul edilmiştir.

Prof.Dr. Nur TUNALI
Vekil Müdür

VARLIK YÖNETİMİ ŞİRKETİNDE SINIFLANDIRMA YÖNTEMİYLE BORÇ GERİ ÖDEME EĞİLİMİ ANALİZİ

ORIJINALLIK RAPORU

% 15 BENZERLIK ENDEKSİ	% 11 İNTERNET KAYNAKLARI	% 3 YAYINLAR	% 11 ÖĞRENCİ ÖDEVLERİ
----------------------------------	---------------------------------------	------------------------	---------------------------------

BİRİNCİL KAYNAKLAR

1	Submitted to Marmara University Öğrenci Ödevi	% 1
2	www.alomaliye.com İnternet Kaynağı	% 1
3	www.academia.edu İnternet Kaynağı	% 1
4	Submitted to Haliç Üniversitesi Öğrenci Ödevi	% 1
5	Submitted to Karadeniz Teknik University Öğrenci Ödevi	<% 1
6	Submitted to Gazi University Öğrenci Ödevi	<% 1
7	journals.manas.edu.kg İnternet Kaynağı	<% 1
8	www.researchgate.net İnternet Kaynağı	<% 1

Handwritten signature

05/07/2019

TEZ ETİK BEYANI

Yüksek Lisans Tezi olarak sunduğum “Varlık Yönetimi Şirketinde Sınıflandırma Yöntemiyle Borç Geri Ödeme Eğilimi Analizi” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof. Dr. Mübariz EMİNLİ'nin sorumluluğunda tamamladığımı, verileri kendim topladığımı, analizleri ilgili laboratuvarlarda yaptığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim.


(İmza)

Can TAŞCI

ÖNSÖZ

Bu çalışma 2014 – 2019 yılları arasında T.C. Haliç Üniversitesi Bilgisayar Mühendisliği Bölümü'nün bilimsel araştırma ve uygulama çalışmalarına verdiği destek ile hazırlanmıştır.

Yüksek lisans eğitimim ve tez çalışmamın tamamlanması süresince büyük bir gayret ve özveriyle çalışmamı takip eden, gösterdiği sabır ve hoşgörüsüyle bana destek olan tez danışmanım Sayın Prof. Dr. Mübariz Eminli'e çok teşekkür ederim.

Tez çalışmam boyunca bana yardımcı olan Sayın Direktörümüz Alper Cumhuri Albayrak, arkadaşım Neslihan Varlı ve ablam Dr. Sibel Taşcı'ya teşekkürü bir borç bilirim.

Son olarak eğitim hayatım boyunca bana destek olan ve verdiğim her kararda yanımda durarak beni bu günlere getiren sevgili anne ve babama, ablam Hülya ve Canan'a sonsuz teşekkür ederim.

İstanbul, 2019

Can TAŞCI

İÇİNDEKİLER

	Sayfa No
TEZ ETİK BEYANI	III
ÖNSÖZ.....	IV
İÇİNDEKİLER.....	I
KISALTMALAR.....	III
ŞEKİLLER	II
ÇİZELGELER	III
ÖZET.....	IV
ABSTRACT	V
1. GİRİŞ.....	1
2. GENEL BİLGİLER.....	3
2.1. Benzer Çalışmalar	3
2.2. Veri Madenciliği	5
2.3. Veri Madenciliği Yöntemleri	7
2.4. Sınıflandırma Teknikleri	8
2.5. Karar Ağaçları.....	10
2.6. Naive Bayes	13
2.7. Nitelik Seçimi	14
2.8. Varlık Yönetimi Şirketi.....	14
3. GEREÇ VE YÖNTEMLER	18
3.1. Veri	18
3.2. Kullanılan Teknolojiler	20
3.2.1. Microsoft Access	20
3.2.2. Weka.....	21
3.3. Veri Toplama	23
3.4. Analiz.....	26
4. BULGULAR.....	37
4.1. C4.5.....	37
4.2. Naive Bayes	49

4.3. C4.5 ve Naive Bayes Bulgularının Karşılaştırılması	50
5. TARTIŞMA.....	52
6. SONUÇLAR.....	55
7. ÖNERİLER.....	59
8. KAYNAKLAR.....	60
9. ÖZGEÇMİŞ	63

KISALTMALAR

BDDK	: Bankacılık Düzenleme ve Denetleme Kurumu
MERNİS	: Merkezî Nüfus İdare Sistemi
UYAP	: Ulusal Yargı Ağı Bilişim Sistemi
v.b.	: Ve benzeriz
VM	: Veri Madenciliği
VY	: Varlık Yönetim

ŞEKİLLER

	Sayfa No
Şekil 2.1. Veri tabanlarında bilgi keşfi süreci ve veri madenciliği	6
Şekil 2.2. Veri madenciliği modelleri ve görevleri.....	7
Şekil 2.3. Veri madenciliği sınıflandırması	8
Şekil 2.4. X ve Y nitelikleri üzerine uygulanan testleri içeren basit bir karar ağacı ...	12
Şekil 3.1. Veri önizlenimi	19
Şekil 3.2. MS ekran görüntüsü	21
Şekil 3.3. Weka uygulaması java gereksinimi	22
Şekil 3.4. Weka giriş ekranı.....	22
Şekil 3.5. Weka keşfet ekranı	23
Şekil 3.6. Borçlu Çizelgesi	24
Şekil 3.7. Borçlu Çizelgesi önizleme.....	25
Şekil 3.8. Adres Çizelgesi.....	25
Şekil 3.9. Borçlu Çizelgesi önizleme.....	26
Şekil 3.10. Weka ile nitelik seçme ekranı.....	27
Şekil 3.11. Niteliklerin sıralanma sonucu.....	28
Şekil 3.12. Karar ağacı.....	38

ÇİZELGELER

	Sayfa No
Çizelge 2.1. Bazı karar ağacı algoritmalarının özellikleri	11
Çizelge 3.1. Niteliklerin alabileceği değerler	19
Çizelge 3.2. Müşteri kategorisi (sınıf) veri dağılımı.....	28
Çizelge 3.3. Borçluların cinsiyet dağılımı	29
Çizelge 3.4. Borçluların yaş dağılımı	29
Çizelge 3.5. Borçluların doğum yeri dağılımı	30
Çizelge 3.6. Borçluların alacaklı sayısı dağılımı	30
Çizelge 3.7. Borçluların anapara dağılımı	31
Çizelge 3.8. Borçluların merkez bankası riski dağılımı	32
Çizelge 3.9. Borçluların ödeme eğilimi dağılımı.....	32
Çizelge 3.10. Borçluların sigorta dağılımı.....	33
Çizelge 3.11. Borçluların toplam bakiye dağılımı.....	33
Çizelge 3.12. Borçluların toplam tahsilat dağılımı	34
Çizelge 3.13. Borçluların ikamet ettiği il dağılımı	34
Çizelge 3.14. Borçluların ikamet ettiği ilçe veri dağılımı	35
Çizelge 3.15. Borçluların ikamet ettiği mahalle dağılım.....	35
Çizelge 4.1. C4.5 sınıflandırma sonucu özeti	38
Çizelge 4.2. C4.5 sınıflara göre detaylı doğruluk.....	46
Çizelge 4.3. C4.5 karışıklık matrisi	47
Çizelge 4.4. C4.5 test verisi karışıklık matrisi.....	48
Çizelge 4.5. Naive Bayes sınıflandırma sonucu özeti	49
Çizelge 4.6. Naive Bayes karışıklık matrisi.....	50

ÖZET

VARLIK YÖNETİMİ ŞİRKETİNDE SINIFLANDIRMA YÖNTEMİYLE BORÇ GERİ ÖDEME EĞİLİMİ ANALİZİ

Varlık Yönetimi şirketleri, bankalardan ihaleyle beraber tahsilatı gecikmiş alacakları devir aldıktan sonra yeniden yapılandırarak tahsil etmeye çalışan şirketlerdir. Bu araştırmada, varlık yönetimi şirketi olarak, borçlunun demografik niteliklerinin (cinsiyet, maaş, yaş, doğum yeri gibi), merkez bankası riski, yapılan tahsilat, bakiye ve profesyonel şirket tarafından oluşturulmuş lokasyonlara göre verilmiş skorlarla birlikte borcun geri ödeme eğilimi analizini tespit etmek amaçlanmıştır.

Araştırmada, farklı kategorilerdeki borçluların geri ödeme eğilimleri veri madenciliği sınıflandırma tekniklerinden C4.5 ve Naive Bayes ile incelenmiş, elde edilen bulgular karşılaştırılmıştır. Borçluların yer aldıkları kategoriler ve borçlu niteliklerin geri ödeme eğilimlerindeki önemi tespit edilmiş böylece Varlık Yönetim şirketlerinin ileriye dönük eylem planı oluşturmasına yardımcı olması hedeflenmiştir.

Anahtar Kelimeler: Varlık Yönetimi Şirketi, Veri Madenciliği, Karar Ağaçları, Naive Bayes, Gecikmiş Alacaklar.

ABSTRACT

ANALYSIS OF DEBT REPAYMENT TREND BY CLASSIFICATION METHOD IN THE ASSETS MANAGEMENT COMPANY

Asset Management companies are the companies that try to collect and restructure the collection after taking overdue collection from the banks. In this research, it is aimed to find the repayment tendency analysis with the borrower's demographic characteristics (such as gender, salary, age, place of birth), central bank risk, collections, balance attributes and location scores by established from professional company.

In this research, repayment trend of borrowers in different categories were compared facts which are founded by data mining classification techniques such as C4.5 and Naive Bayes. The importance of the categories of borrowers and the characteristics of the borrowers in the repayment tendencies were determined and thus, it was aimed to help Asset Management companies to discover a way for the future plan.

Keywords: Asset Management Company, Data Mining, Decision Trees, Naive Bayes, Overdue Collection

1. GİRİŞ

Türk bankacılık sisteminin yeniden yapılanmasında önemli bir yeri olan 4389 sayılı Bankalar Kanunu ile bankacılık sektörünü hem tek elden denetleyecek denetim mekanizmaları oluşturulmuş hem de uluslararası standartlar çerçevesinde yönetilmesi sağlanmıştır. Bu nedenle oluşturulan kredi kayıp karşılıkları sürecin denetlenmesinde ve kontrol altına alınmasında önemli bir yere sahiptir. Bankacılık faaliyetlerinin sürdürülmesinde kayıpların en aza indirilebilmesi için kredi kayıp karşılıkları ayırmak gerekmektedir. Bu bakımdan kredi kayıp karşılıklarının bankacılığın sigortası olduğu söylenebilir (Özgür, 2000: 32).

Günümüzde hayatın her alanında yaşanan gelişmeler ile birlikte insanların ihtiyaçları ve yaşam standartları da değişmektedir. Bu nedenle zaman zaman insanlar, gerek yaşam standartlarını devam ettirebilmek gerekse zorunlu ihtiyaçlarını karşılayabilmek için bankalardan kredi almak durumunda kalabilmektedirler. Alınan kredilerin zamanında geri ödenememesi durumunda bankalar, borçlular hakkında yasal süreç başlatarak borcun tahsil etme yoluna gitmektedir. Bu sürecin hem çok masraflı olması hem de geri dönüş olasılığının düşük olmasından dolayı bankalar bu duruma ayrıca bütçe ayırmak istememektedir. Geri dönüş olasılığı düşük olan krediler batık kredi olarak adlandırılmakta olup batık krediye sahip olan bankalar bu krediyi geri dönmeyecek bir yatırım olarak değerlendirmekte ve anapara üzerinden bir paket hazırlayıp ihale usulüyle varlık yönetimi şirketlerine sunmaktadırlar. Bu süreçte varlık yönetimi şirketleri de bu paket üzerinde gerekli çalışmaları yaptıktan sonra bankaya kendi tekliflerini vermekte ve en iyi teklifi veren varlık yönetimi şirketi tüm haklara sahip olmaktadır.

Varlık Yönetimi şirketleri, bankalardan tahsilatı gecikmiş alacakları devir alarak yeniden yapılandırarak tahsil etmeye çalışan şirketler olup finans özelliklerinden dolayı BDDK'a bağlı olarak çalışmaktadırlar. Bankalar devir işlemi tamamladıktan sonra tüm yasal haklara sahip olan varlık yönetimi şirketi borcu yeniden yapılandırıp tahsil etmeye çalışmaktadır.

Çalışmanın amacı varlık yönetimi şirketi olarak, borçlunun demografik niteliklerin (cinsiyet, maaş, yaş, doğum yeri gibi), merkez bankası riski, yapılan tahsilat, bakiye ve profesyonel şirket tarafından oluşturulmuş lokasyonlara göre verilmiş skorlarla birlikte borcun geri ödeme eğilimi analizini tespit etmektir. Elde edilen bulgular sonucunda varlık yönetimi şirketlerinin ileriye dönük eylem oluşturabilmesinde yardımcı olması hedeflenmektedir. Araştırmanın amacına yönelik olarak borcun geri ödenmesindeki yeni süreçlerin bulunması veri madenciliği kullanılarak yapılabilir.

Birinci bölümde 4389 sayılı Bankalar Kanunu doğrultusunda gerçekleştirilen bankacılık faaliyetlerine ve kayıp kredilere giriş yapılarak varlık yönetim şirketleri açısından kayıp kredilerin önemine yer verilmiştir. Ardından da varlık yönetim şirketleri açısından batık kredilerin geri ödenme eğilimlerinin analizinin hedeflendiği çalışmanın amacına yer verilmiştir.

2. GENEL BİLGİLER

2.1. Benzer Çalışmalar

Veri madenciliği, varlık yönetim sistemleri, batık krediler, kredi kartı dolandırıcılıkları, müşteri memnuniyet arařtırmaları, sosyal medya arařtırmaları gibi birçok sektörde farklı ya da benzer amaçlar ile kullanılmaktadır (Demir, 2018). Günümüze kadar veri arkeolojisi, veri taranması, veri balıkçılığı, özbilgi çıkarımı, veri tabanlarında özbilgi keşfi gibi birçok farklı şekilde adlandırılan veri madenciliği, farklı isimler altında istatistik ve başka birçok alanda kullanılmıştır (Yaşar, 2017).

Günümüzde teknolojiye yaşanan hızlı deęişimler toplumların yaşayışlarını da etkilemekte ve deęişimler sonucunda bireylerin aldıkları kararlarda önsezilere dayalı yanlış karar alma riski yüksek olabilmektedir. Riski azaltmanın en önemli yolu bilgiye dayalı yönetimi öngören karar destek çözümleridir (Savaş, Topalođlu ve Yılmaz, 2012). Bu bakımdan karar destek sistemi oluřturmada veri madenciliği yöntemleri önemli birer araç olarak karřımıza çıkmaktadır. Borcun geri ödenme eğilimlerinin incelenbilmesi için veri madenciliğinde kullanılan sınıflandırma yöntemlerinden karar ağaçları ve Naive Bayes algoritmaları kullanılacak, elde edilen sonuçlar karřılařtırılarak hangi algoritmanın daha verimli olduđu belirlenmeye çalışılacaktır. İlgili literatür incelendiğinde veri madenciliğine yönelik olarak bir çok arařtırma olduđu görülmüştür.

Çakır (2008) tarafından yapılan ve veri madenciliğinin sınıflandırma fonksiyonuna ilişkin tekniklerini gerçek yaşam verileri üzerinde uygulamak, veri madenciliği sürecini tüm aşamaları ile gerçekleřtirmek, uygulamada elde edilen sonuçlarla tekniklere ilişkin farklılıkları tartıřmak ve hangi tekniğin hangi kořullarda uygulanmasının daha etkin olacađına yönelik önerilerde bulunmasının amaçlandığı çalışmada veri sayısının ve veri kalitesinin, uygulamaların başarısında önemli birer faktör olduđu sonucuna ulařılmıştır.

Ceylan (2006) tarafından yapılan Türkiye'deki batık kredi kart borç sahiplerinin profili belirlenmeye çalışılmış olan çalışmada tüketici kredisinin gelir seviyesi, ev sahibi olup olmaması ve cinsiyetiyle tüketici kredisinin geri dönmemesi arasında anlamlı bir ilişki olduğu belirtilmiş iken incelenen diğer değişkenlerden tüketici kredisinin yaşı, medeni durumu, öğrenim durumu, gelirinin kaynağı, kredi kartı sahibi olup olmaması ile kullanılan tüketici kredisinin mahiyeti, miktarı ve vadesi arasında anlamlı bir ilgiye rastlanılmadığı sonucuna ulaşılmıştır.

Hörkkö (2010) tarafından yapılan Finlandiya'da tüketici kredisinin alınırken sosyodemografik ve davranışsal özelliklerin etkisi araştırıldığı çalışmada sosyodemografik ve davranışsal özelliklerin dikkate değer bir etki yarattığı ve bunların en önemli sosyodemografik nitelikleri gelir, yaş, eğitim ve milliyet olurken kredi kullanılabilirlik skoru, borcun büyüklüğü gibi davranışsal özelliklerinden önemli bir güce sahip olduğu sonucuna ulaşılmıştır.

Yılmaz ve ark. (2013) tarafından yapılan tüketicilerin yaş, cinsiyet ve medeni durum gibi demografik değişkenlerle kredi kartlarını ilgilendiren birçok kategorik değişken arasında anlamlı ilişkiler ortaya çıkarmaya çalışılan ve tüketicileri bazı kategorilere girmeye daha meyilli oldukları gözlemlenen çalışmada yaş, cinsiyet, medeni durum, pozisyon ve gelir düzeyi gibi demografik türde birçok değişkenle kredi kartını ilgilendiren kategorik değişkenler arasında çok sayıda anlamlı ilişkiler olduğunu sonucuna ulaşılmıştır.

Amponsah (2016) tarafından yapılan öğrencilerin borçlu olmalarını, borcunu geri ödemelerini incelemek ve öğrenci borçlularının geri ödemelerinde cinsiyet, yaş ve tutumlarına etkilerini araştırdığı çalışmada öğrencilerin borçlularının cinsiyetlerinin borç geri ödemelerini olumlu yönde etkilediği tespit edilmiş ve kız öğrenci borçlularının, erkek meslektaşlarına göre biraz daha fazla seviyede borcunu geri ödedikleri sonucuna ulaşılmıştır.

Ramanujam and Vidya (2017) tarafından yapılan çalışmada ise demografik özelliklerin Hindistan'ın Virudhunagar ilçesindeki bir kuruluş için kredi geri ödeme performansına etkileri araştırılmıştır. Araştırma sonucunda cinsiyet, yaş, eğitim düzeyi, mesleki geçmiş deneyim yılı, gibi demografik özellikler ile kredi geri ödeme performansları arasında anlamlı bir fark oluşturmadığı sonucuna ulaşılmıştır.

Bilen (2009) tarafından yapılan veri madenciliği yöntemlerinden sınıflandırma ve kümeleme ile etkili bir personel seçim mekanizması geliştirilerek özellikle personel

seçimi sürecinde fayda sağlanması amaçlandığı çalışmada doğru insanın doğru özellikler ile doğru yerde kullanılması ve doğru hedeflere odaklanması ile ciddi anlamda bir fayda sağlanacak, gerek çalışanlar gerekse de işletme için verimlilik ve etkinliğin artacağı sonucuna ulaşılmıştır. Bu araştırmanın sonuçları, benzer şekilde müşteri seçiminde ve sınıflandırılmasında bankalar gibi kredi veren sektörler tarafından geliştirilecek ve veri madenciliğinin etkili bir şekilde kullanılacağı bir sistemin oluşturulmasında kullanılabileceği şeklinde yorumlanabilir.

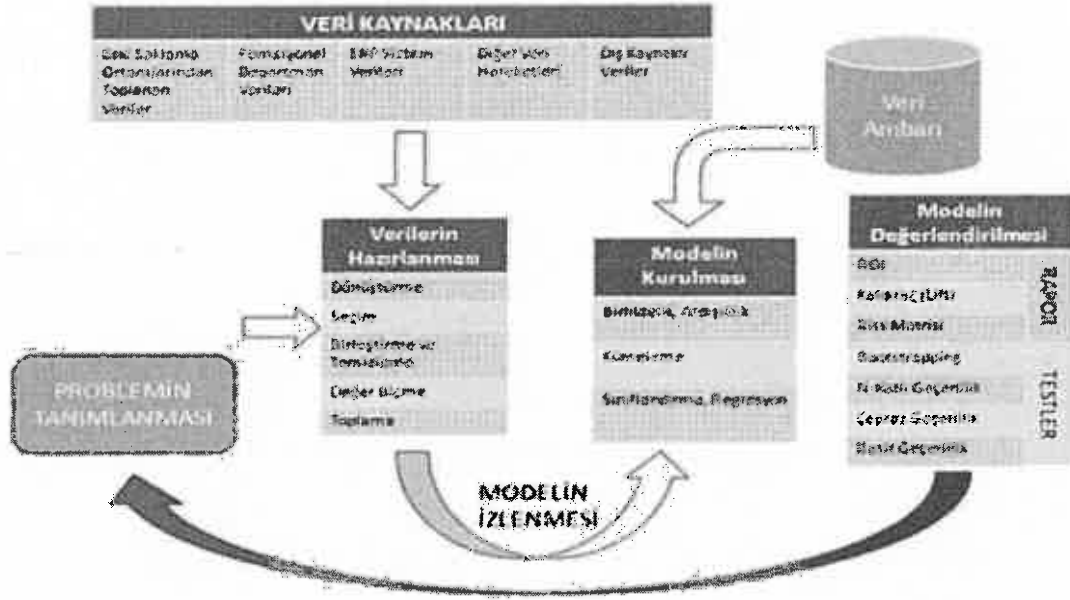
Ancak varlık yönetim sistemlerinin borçlu eğilimlerinin incelenmesine yönelik olarak yeterli sayıda araştırma olmadığı görülmüştür. Bu araştırma ile veri madenciliği sınıflandırma teknikleri kullanılarak varlık yönetim sistemlerinin borçlunun borcu geri ödeme eğilimlerini değerlendirilmesine yönelik önemli bir katkı sağlaması hedeflenmektedir.

2.2. Veri Madenciliği

1980'li yılların sonunda ortaya çıkan veri madenciliği 1990'lı yıllarda büyük adımlar atmış ve günümüzde gelişmeye devam etmekte olup multidisipliner bir alandır. Veri madenciliği büyük miktarda veriden veri çıkarma sürecidir (Han and Kamber, 2001).

Veri depolarında yer alan keşfedilmemiş bilgilerin, veriler arasındaki kavramsal ve niteliksel ilişkilerin, ayrışma ve yönelimlerin çeşitli modelleme ve analizler yardımı ile mantıksal tahminlemeler yapılarak çıkarılması olarak da tanımlanabilen veri madenciliğine yönelik olarak literatürde birçok farklı tanım olduğu görülmektedir (Alagöz ve ark., 2014).

Alanındaki en önemli uzmanlardan olan Piatetsky-Shapiro'ya göre veri madenciliği; kümeleme, veri özetleme, sınıflama kurallarının öğrenilmesi, bağımlılık ağlarının bulunması, değişikliklerin analizi ve anomali tespiti gibi birçok teknik yaklaşım içermekte olup eldeki verilerden daha önceden bilinmeyen ancak yararlı olan bilgilerin monoton olmayan bir süreçte çıkartılması işlemidir. Veri madenciliği süreci Şekil 1'de verilmiştir (Akpınar, 2000).



Şekil 2.1. Veri tabanlarında bilgi keşfi süreci ve veri madenciliği (Akpınar, 2000)

Hand et al. (2001) veri madenciliğini, beklenmeyen ilişkileri bulmak ve verileri veri sahibi için hem anlaşılabilir hem de kullanışlı olan yeni yollarla özetlemek için (genellikle büyük) gözlemsel veri setlerinin analizi olarak tanımlamaktadır.

Kalikov'a (2006) göre ise veri madenciliği büyük miktarlardaki verinin içinden geleceğin tahmin edilmesinde yardımcı olabilecek yararlı bağlantıların ve kuralların bilgisayar programlarının aracılığıyla aranması ve analizidir. Veri madenciliğinin en önemli özelliklerinden biri veri grupları arasındaki benzer eğilimlerin belirlenmesini sağlaması iken bir diğeri ise daha önceden bilinmeyen, veri ambarları içinde bulunmasına karşın ilk etapta görülemeyen bilgileri ortaya çıkarabilmesidir (Savaş ve ark., 2012). Yani veri madenciliği araçları yardımıyla şirketler, destek sistemlerinde gerekli olan eğilimleri ve davranış kalıplarını ortaya çıkararak daha etkin kararlar alabilmektedirler (İnan, 2003).

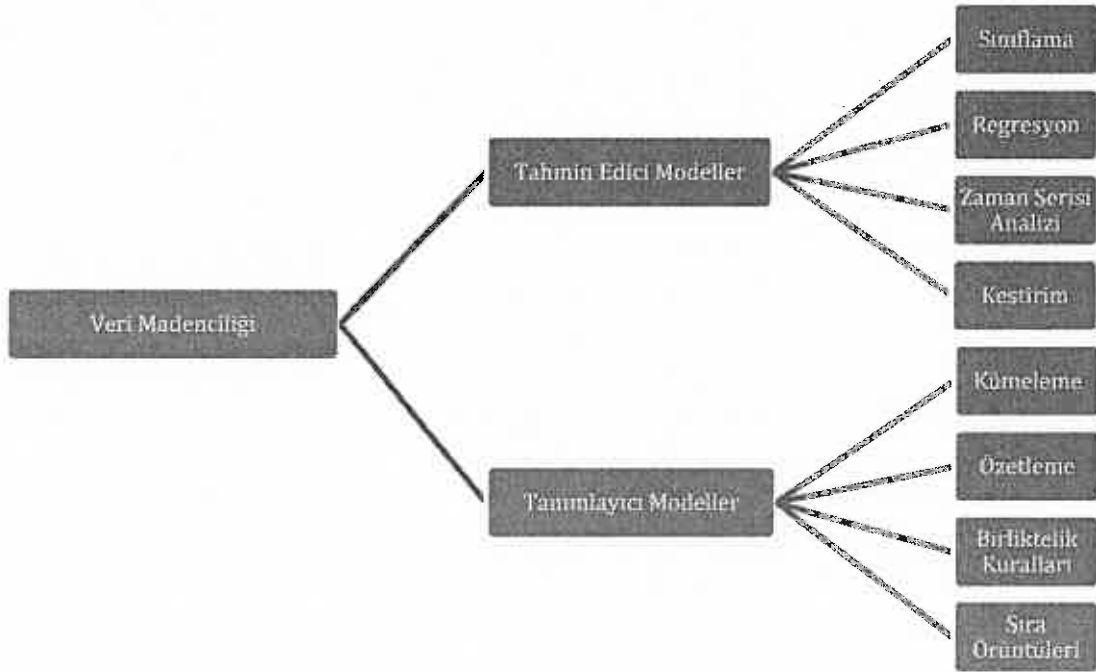
Görüldüğü üzere veri madenciliği günümüzde birçok alanda kullanılmakta olup farklı tanımlarla karşımıza çıkabilmektedir. Tüm tanımların ortak yönleri incelendiğinde ise veri madenciliğinin veri ambarları içinde bulunan verilerin arasındaki ilişkilerin analizi olduğu söylenebilir.

2.3. Veri Madenciliği Yöntemleri

Temel olarak veri madenciliği, yazılım tekniklerindeki teknolojiyi kullanarak verilerin analiz edilmesidir. Farklı görevleri yerine getirmek için pek çok farklı algoritmanın kullanıldığı veri madenciliğinde bu algoritmalar verilere uygun modeli bulmaya çalışarak verileri incelemekte ve özelliklerine en uygun modeli seçmektedir (Aydın ve Özkul, 2015). Veri madenciliği modellerinin sınıflandırılmasına ilişkin literatürde farklı sınıflandırmalar karşımıza çıkmaktadır.

Veri madenciliği, veritabanı sistemleri, istatistikler, makine öğrenmesi, görselleştirme ve bilgi bilimi de dahil olmak üzere bir dizi disiplinin birleştiği disiplinlerarası bir alandır. Veri madenciliğine katkıda bulunan disiplinlerin çeşitliliği nedeniyle, veri madenciliği araştırmalarının çok çeşitli veri madenciliği sistemleri üretmesi beklenmektedir. Bu nedenle, veri madenciliği sistemlerinin net bir şekilde sınıflandırılmasını sağlamak gerekmektedir (Han and Kamber, 2000).

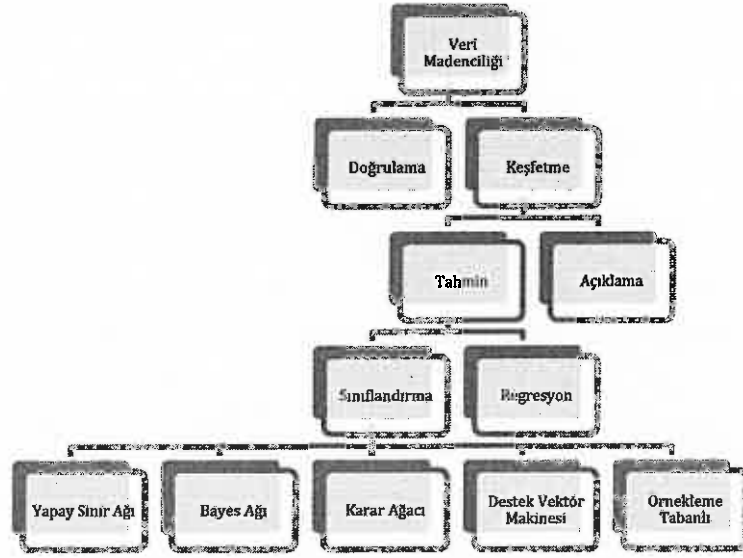
Dunham'a (2003) göre veri madenciliği "tahmin edici" ve "tanımlayıcı" modeller olmak üzere iki kategoriye ayrılmaktadır. Bu modeller Şekil 1.2'de gösterilmektedir.



Şekil 2.2. Veri madenciliği modelleri ve görevleri (Dunham, 2003)

Veri madenciliği yöntemleri Şekil 2.3'de görüldüğü üzere doğrulama ve keşfetme olarak 2 gruba ayrılmaktadır. Doğrulama yöntemleri bir dış kaynak tarafından önerilen bir hipotezin değerlendirilmesi ile ilgilidir. Keşfetme yöntemleri verilerdeki kalıpları tanımlayan yöntemler olup açıklama ve tahmin yöntemlerinden

oluşmaktadır. Tahmin odaklı yöntemler davranışsal bir model oluşturmayı amaçlamaktadır ve bir veya birden fazla değişkenin değerlerini tahmin edebilmektedir (Maimon and Rokach, 2010).



Şekil 2.3. Veri madenciliği sınıflandırması (Maimon ve Rokach, 2010)

İlgili literatür incelendiğinde veri madenciliğinin mevcut veri tabanından tahminler yapmak (tahmin edici model) ve mevcut veriden çıkarılabilecek davranışları betimlemek (tanımlayıcı bilgi) olmak üzere iki önemli soruna cevap bulmayı amaçladığı söylenebilir. Bu bakımdan tahmin edici modeller genellikle verilerdeki eğilimleri ve ilişkileri bulmak için kullanılırken, tanımlayıcı modeller ise örüntü ve ilişkileri bulmada veriyi görselleştirmede kullanıldığı söylenebilir (Yaşar, 2017). Bu bakımdan ileri vadeli borçların başarısıyla ilgili olarak bankacılık sektörü, dolandırıcılık olasılıklarıyla ilgili olarak sigorta sektörü tahmin edici modeller için örnek olarak verilebilir (Birtül, 2011).

Sonuç olarak veri madenciliği modellerini, gördükleri işlemlere göre sınıflama ve regresyon, kümeleme, birliktelik kuralları olmak üzere temel olarak 3 grupta toplamak mümkündür (Savaş ve ark., 2012).

2.4. Sınıflandırma Teknikleri

Sınıflandırma ve regresyon, yeni bir kaydın özelliklerini esas alarak daha önceden tanımlanmış sınıflardan uygun olanına atamaktır. Yani veri madenciliğinde sınıflandırması yapılacak olan her bir kayıta belirli bir sınıf etiketi atanmaktadır (Deveci, 2018). Burada önemli olan her bir sınıfın özelliklerinin daha önceden açıkça belirlenmiş olmasıdır. Sınıflandırmada bir değişken bağımlı değişken olarak atanırken

diğerleri bağımsız deęişkenler olarak atanmaktadır. Sınıflandırmada amaç bağımsız deęişkenlerin yer aldığı bir modelde bağımlı deęer açısından anlamlı bir yöntem kurmaktır (Akçay, 2014). Bu açıdan veri sınıflandırma, mevcut verilerin daha etkili ve verimli olarak kullanılabilmesi için verileri kategorilere göre düzenleme olup verilerin bulunmasını kolaylaştırdığı söylenebilir.

Sınıflandırma tekniğinde ilk işlem olan veri setindeki nesnelere ifade eden bir sınıf deęişkeni atanmasının ardından öğrenme kümesi adıyla bir model kurulur ve her sınıfı temsil eden öğrenme kümesi nesnelere belirlenir. Öğrenme kümesi dışındaki nesnelere de deneme kümesi (test set) olarak adlandırılır. Son olarak deneme kümesi, nesnelere modele uygun sınıflandırılır. Öğrenme ve test kümelerinin belirlenmesinde farklı yöntemler kullanılabilmekte olup kullanılan yöntem, kurulan modelin başarısını önemli ölçüde etkilemektedir (Erduran, 2017).

Veri madencilięi modellerinden biri olan sınıflandırma ve regresyon, önemli veri sınıflarını ortaya koymada veya gelecek veri eğilimlerini tahmin etme modellerini kurmada kullanılan analiz yöntemleri olup sınıflandırma, kategorik deęerleri tahmin etmede kullanılırken regresyon ise süreklilik gösteren deęerlerin tahmin edilmesinde kullanılmaktadır. Başlıca sınıflama ve regresyon teknikleri şunlardır (Çalış ve ark., 2014):

1. Yapay Sinir Ağları (Artificial Neural Networks)
2. Genetik Algoritmalar (Genetic Algorithms)
3. K-En Yakın Komşu (K-Nearest Neighbour)
4. Naive-Bayes sınıflayıcısı
5. Lojistik Regresyon
6. Karar Ağaçları (Decision Trees)

Sonuç olarak veri sınıflandırmanın, verileri kolayca aranabilir ve izlenebilir hale getirmek için etiketlenmesini sağlamanın yanı sıra arama işlemini hızlandırabilen, depolama ve yedekleme maliyetlerini ise düşürebilen birden fazla veri yinelenmesini de ortadan kaldırdığı söylenebilir. Bu açıdan bakıldığında veri tabanı sistemlerine kolay entegre olması ve kolay yorumlanması sebebiyle en çok kullanılan sınıflandırma yöntemleri; karar ağaçları ve mesafeye dayalı algoritmalarıdır (Erduran, 2017). Diğer taraftan verilerin sınıflandırılmasında kullanılan bir başka yöntem olan Naive Bayes ise verilerin hangi sınıfa ait olduğunu belirlemede olasılıksal hesaba dayanan bir algoritmadır (Arpacı ve Kalıpsız, 2018).

2.5. Karar Ağaçları

Veri madenciliği analizlerinde kullanılan yöntemlerden biri de karar ağaçlarıdır. Karar ağaçlarında amaç sınıf etiketleri bilinen veri örneklerinden yola çıkarak bir ağaç yapısı oluşturmaktır. Ağaç yapısının en üst seviyesinde yer alan bileşenine kök, en altta yani uçlardaki bileşenlerine yaprak ve tüm bu bileşenleri birbirine bağlayan yapılara ise dal adı verilmektedir. Karar ağacı algoritmasının çözmesi gereken en temel problem, en üst seviyede yer alan kökte hangi özneliğin yer alacağını ve düğümler ve yapraklar oluşturulurken kökten itibaren hangi kritere göre hareket edileceğinin tespit edilmesidir (Güldal ve Çakıcı, 2017). Bir karar ağacı, her düğümün bir nitelik değeri üzerinde bir testi ifade ettiği, her dalın bir test sonucunu temsil ettiği ve ağaç yapraklarının sınıfları veya sınıf dağılımlarını temsil ettiği akış şeması benzeri bir ağaç yapısıdır ve karar ağaçları kolayca sınıflandırma kurallarına dönüştürülebilmektedirler (Han and Kamber, 2000).

Düğüm, dal ve yapraklardan oluşan karar ağaçlarında karar düğümü gerçekleştirilecek test olup testin sonucunda ağaç dallara ayrılmaktadır. Ağacın her bir dalı sınıflama işlemini tamamlayabilmektedir. Bir dalın ucunda sınıflama gerçekleşmemesi durumunda bir karar düğümü, sınıflama gerçekleşmesi durumunda ise dalın sonunda yaprak oluşmaktadır ve bu yaprak veri üzerinde belirlenmek istenen sınıflardan birisidir (Gökgöz, 2015).

Karar ağacı, kök ve her iç düğümün soru ile etiklendiği bir ağaç olup her bir düğümden çıkan dallar, ilişkili soruya verilen her bir cevabı, her bir yaprak düğümü ise söz konusu probleme bir çözüm öngörüsünü temsil etmektedir. Karar ağacı yüksek boyutlu verileri işleyebilmekte olup karar ağacı sınıflandırıcısının en önemli özelliği, karmaşık bir karar verme sürecini daha basit kararların toplanmasına bölme ve yorumlanması daha kolay olan bir çözüm sağlama yeteneğidir (Seema et al., 2012).

Sonuç olarak karar ağacı ile ilgili bölme kriterlerine göre yapılan sınıflandırma tekniklerinden biri olup örnekleri özellik (özellik) değerine göre sıralayarak sınıflandıran bir ağaç yapısı gibi bir akış şemasıdır. Bir karar ağacındaki her düğüm, sınıflandırılacak bir durumda bir özelliği temsil etmekte ve tüm dallar testin sonucunu belirttiği söylenebilir (Sharma et al., 2013). Tahmin edici ve tanımlayıcı özelliklere sahip olan karar ağaçları kuruluşlarının ucuz olması, kolayca yorumlanabilmeleri, veri tabanlarına kolayca entegre edilebilmeleri ve güvenilir olmalarından dolayı veri madenciliğinde sıklıkla kullanılmaktadır (Akpınar, 2000). Karar ağaçları yönteminde

CHAID, QUEST, CRT, SLIQ, ID3, C4.5 ve C5.0 gibi algoritmalar kullanılmaktadır. Bu karar ağaçlarının özellikleri Çizelge 2.1’de verilmiştir.

Çizelge 2.1. Bazı karar ağacı algoritmalarının özellikleri (Ersöz, Merdin ve Ersöz, 2015)

Karar Ağacı Algoritması	Özellikler
CRT	Gini’ye dayalı ikili bölme işlemi mevcuttur. Son veya uç olmayan her bir düğümde iki adet dal bulunmaktadır. Budama işlemi ağacın karmaşıklık ölçüsüne dayanır. Sınıflandırma ve regresyonu destekleyici bir yapıdadır. Sürekli hedef değişkenleri ile çalışır. Verinin hazırlanmasına gereksinim duyar.
C4.5 ve C5.0	Her düğümden çıkan çoklu dallar ile ağaç oluşturur. Dalların sayısı tahmin edicinin kategori sayısına eşittir. Tek bir sınıflayıcıda birden çok karar ağacını birleştirir. Ayırma işlemi için bilgi kazancı kullanır. Budama işlemi her yapraktaki hata oranına dayanır.
CHAID	Ki-kare testleri kullanarak bölme işlemi gerçekleştirir. Dalların sayısı iki ile tahmin edicinin kategori sayısı arasında değişir.
SLIQ	Hızlı ölçeklenebilir bir sınıflayıcıdır. Hızlı ağaç budama algoritması mevcuttur.
SPRINT	Büyük veri kümeleri için idealdir. Bölme işlemi tek bir niteliğin değerine dayanır. Tüm bellek sınırlamaları üzerinde nitelik listesi veri yapısı kullanarak işlem yapar.

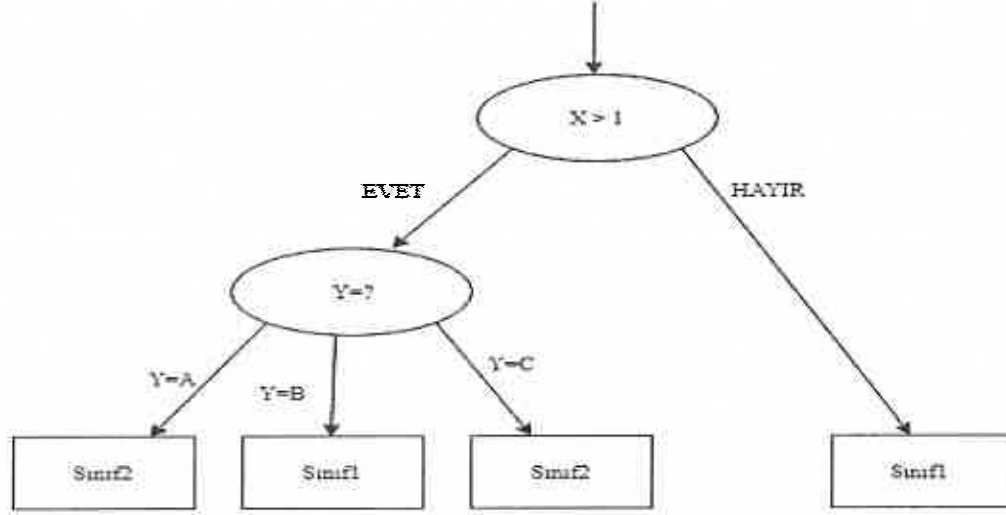
C4.5 algoritması belirli bir veri örneğine (tek değişkenli veya çok değişkenli tahmin ediciler) dayanarak, veri madenciliğinde bir karar vermede karar ağacı sınıflandırıcısı olarak kullanılmaktadır. C4.5, bilgi entropi kavramını kullanarak karar ağaçlarını ID3 ile aynı şekilde bir dizi eğitim verisinden oluşturmaktadır.

Sınıflandırma işlemini verilerin sıklıklarına göre yapmakta olan bir algoritma olup verilerdeki belirsizliğin ölçüsüne entropi hesabı ile yapılan C4.5 algoritmasında; veri kümesinde yer alan tüm sütunların $-\sum_{i=1}^n p_i \log_2 p_i$ formülüne göre entropisi hesaplandıktan sonra her bir sütunun kazancının hesaplanması için veri kümesindeki her bir sütunun entropisi, sınıfın entropisine bölünür ardından kazancı en yüksek olan kök düğüm olarak atanır (Aksu ve Karaman, 2017). Bu durumda C4.5’in uygulama adımları şu şekilde ifade edilebilir (Agrawal and Gupta, 2013):

1. Taban bilgileri kontrol edilir.
2. Her özellik için bir hesaplama yapılır (Her bir niteliğin bilgi kazanç değeri hesaplanır)
3. En yüksek bilgi kazancına sahip en iyi nitelik seçilir.
4. Seçilen en iyi nitelik ana düğüm olarak ifade edilir.
5. Bu durum bütün alt düğümler oluşana kadar tekrarlanır.

X ve Y’den oluşan iki giriş niteliğine sahip bir örnek sınıfın basit karar ağacı Şekil 1.4’de verilmiştir. Şekil 1.4 incelendiğinde; $X > 1$ olması durumunda $Y = B$

değerlerini taşıyan örneklerin sınıf 1'de, $Y=A$ ve $Y=C$ koşullarına uygun olanların sınıf 2'de; Y 'nin değerini gözönüne almadan $X \leq 1$ koşuluna uygun örneklerin ise sınıf 1'de yer aldığı görülmektedir.



Şekil 2.4. X ve Y nitelikleri üzerine uygulanan testleri içeren basit bir karar ağacı (Özkan, 2008)

Karar ağacı sonuçlarını doğru-pozitif oranı (true-positive rate), yanlış-pozitif oranı (false-positive rate), duyarlılık (precision), kesinlik (recall) ve F_1 skoruna göre değerlendirilebilmektedir. Sınıflandırma doğruluğu, her zaman negatif sınıflandırıcıya aşırı değer verdiği için bu gibi durumlarda makul bir değerlendirme ölçütü değildir. Ayrıca sınıf dengesizliğini doğruluk yardımıyla maliyete duyarlı versiyonları ile ayarlamak da doğru değildir, çünkü bu sadece gerçek negatiflerin faydasını değil, aynı zamanda yanlış pozitiflerin maliyetini de küçümseyecektir. Bu durumda iyi bir çözüm, gerçek negatifleri bir bütün olarak görmezden gelmek ve yanlış pozitif oran yerine performans ölçütü olarak pozitif tahminler arasındaki doğru pozitiflerin oranı olarak tanımlanan hassasiyeti kullanmaktır (Flach ve Kull, 2015). Doğru pozitif, yanlış pozitif, doğru negatif ve yanlış negatif gibi değerlerin yardımıyla duyarlılık (precision), kesinlik (recall), f_1 skoru (F_1 -score veya F-measure) formülleri Denklem (2.1), (2.2) ve (2.3)'de gösterilmektedir:

$$\text{duyarlılık} = \frac{\text{doğru pozitif}}{\text{doğru pozitif} + \text{yanlış negatif}} \quad (2.1)$$

$$\text{kesinlik} = \frac{\text{doğru pozitif}}{\text{doğru pozitif} + \text{yanlış pozitif}} \quad (2.2)$$

$$f1 \text{ skoru} = \frac{2 * \text{kesinlik} * \text{duyarlılık}}{\text{kesinlik} + \text{duyarlılık}} \quad (2.3)$$

(2.1) denkleminde verilen duyarlılık, pozitif olarak tahmin edilen bir durumdaki başarıyı gösteren durumu, (2.2) formülünde verilen kesinlik, pozitif durumların ne kadar başarılı tahmin edildiğini ve (2.3) formülünde verilen F_1 skoru ise kesinlik ve duyarlılığın harmonik ortalamasını göstermektedir.

Büyük veri yığınları içerisindeki gizli ilişki ve örüntülerin farklı araç ve gereçleri kullanarak ortaya çıkarılması şeklinde tanımlanan veri madenciliğinde doğru sonuca ulaşabilmek açısından model kurulumu öncesinde kullanılacak verinin niteliklerinin seçimi büyük önem arz etmektedir (Budak, 2018).

2.6. Naive Bayes

İsmini Bayes Teoremini öneren Thomas Bayes'ten alan, veri madenciliği için en etkili algoritmalarından biri olan ve tüm nitelikleri sınıf değişkenine göre bağımsız bir şekilde değerlendiren bir algoritmadır (Zhang, 2004; Vikramkumar et al, 2014). Veride yer alan kriterlerin sonuca yaptıkları etkileri olasılıksal olarak hesaplayan, Bayes teoremine dayanmakta olan Naive Bayes sınıflandırma yönteminde; $j \in \{1, \dots, J\}$ olmak üzere y_j J tane sınıftan birini temsil eden ayrık bir değişken ve X özelliği ise m adet özellikten oluşan $X=(x_1, x_2, \dots, x_m)$ özellik vektörü olmak üzere kullanılan kurallar denklem 2.4, 2.5, 2.6 ve 2.7'de verilmiştir (Arpacı ve Kalıpsız, 2018):

Olası y_j değeri için Bayes teoremine göre sonsal olasılık $p(y_j | X)$ olmak üzere;

$$P(y_j | X) = \frac{P(X|y_j)P(y_j)}{P(X)} \quad (2.4)$$

Y' hedef sınıfının tahmini için tüm özelliklerin koşullu olasılıklarının çarpımı yapıldığında;

$$Y' \leftarrow \operatorname{argmax}_{y_j} \frac{P(y_j) \prod_{i=1}^m P(X=x_i|y_j)}{\sum_{j=1}^J P(y_j) \prod_{i=1}^m P(X=x_i|y_j)} \quad (2.5)$$

eşitliği elde edilmektedir. Bu eşitliğin sadeleştirilmesi istendiğinde ise;

$$Y' \leftarrow \operatorname{argmax}_{y_j} P(Y = y_j) \prod_{i=1}^m P(X = x_i | Y = y_j) \quad (2.6)$$

$$P(y_j) = \frac{|N_j|}{|N|} \quad (2.7)$$

eşitlikleri elde edilmektedir.

Denklem 2.7. incelendiğinde $|N|$ eğitim niteliklerinin toplam sayısını, $|N_j|$ ise y_j sınıfındaki eğitim niteliklerinin sayısını ifade etmektedir. Böylece Naive Bayes sınıflandırmasında Denklem 2.6. ile her bir sınıfın olasılık değerleri hesaplanmakta ve hesaplama işleminin ardından en büyük olasılığa sahip olan sınıf hedef sınıf olarak seçilmektedir.

2.7. Nitelik Seçimi

Veri madenciliğinde karar ağacının verimli ve tutarlı bir şekilde oluşturulmasında nitelik seçiminin önemi göz önüne alındığında; bazı problemler için tüm özellikler önemli olabilirken, bazı hedef konseptler için sadece küçük bir özellik daha verimli sonuçlar çıkarılmasında yardımcı olabilmektedir.

Veri madenciliği algoritmasını hızlandırmak, veri kalitesini ve veri madenciliğinin performansını arttırmak ve madencilik sonuçlarının anlaşılabilirliğini arttırmak açısından nitelik seçimi yapılırken; nitelik alanının boyutu azaltılmakta, gereksiz, alakasız veya gürültülü veriler silinmektedir. Nitelik seçimi sırasında en çok kullanılan algoritmalarından biri olan Bilgi Kazanç Nitelik Değerlendirme (Information Gain Attributes Evaluation) algoritması, alakasız veya gereksiz nitelikleri tespit edilmesinde yararlı olmaktadır.

Araştırmanın amacına yönelik olarak borcun geri ödenmesindeki yeni süreçlerin bulunmasında veri madenciliği yöntemlerinden biri olan C4.5 karar ağacı yönteminin kullanılacağı bu çalışmada Varlık Yönetimi Şirketlerinin borçluların borcunu geri ödeme eğilimleri incelenecektir.

2.8. Varlık Yönetimi Şirketi

Ülkelerin ekonomilerinin en önemli bileşenlerinden biri de bankacılık sistemidir. Ulusal ve uluslararası alanlarda finansal açıdan birçok işlemi gerçekleştiren bankalar müşterilerine havale, senet ve çek tahsili, kredi gibi farklı hizmetler sunmaktadırlar. Gerçek veya tüzel kişilere verilen ödünç para karşılığında oluşan borcun ödenmesi olarak adlandırılan kredilerin ülkelerin ekonomisinde önemli bir payının olduğu tartışılmaz bir gerçektir. Vadelerine, amaçlarına ve niteliklerine göre birçok çeşidi bulunan kredilerin bankalara geri dönüşünde riskler oluşabilmektedir. Kredilendirme sonucunda ödeme sıkıntısı yaşayan müşteriler ödeme taahhütlerini

yerine getirememesi durumunda bankacılığın likidite yapısında problemler ortaya çıkabilmektedir ve bu durum kredi riski olarak adlandırılmaktadır (Ture, 2015).

Bankalar para ve sermaye piyasalarında çok farklı risklere maruz olmalarından dolayı risk yönetimi konusunda farklı yöntemler geliştirmek durumunda kalmaktadırlar. Bu nedenle bankalarda kredilendirme süreci öncesinde yanlış teminatlandırma, kredi tür ve tutarının yanlış belirlenmesi gibi birçok nedenle kredinin başlangıçtan itibaren sorunlu olmasına neden olabilmektedir (Atalar ve Şahin, 2013). Öte yandan bankalar daha büyük karlara daha büyük risk maliyetlerine katlandığında ulaşabilmektedirler. Banka aktiflerini oluşturan bileşenlerden birisi de kısa vadede nakde dönüştürülemeyen kredilerdir (Yıldırım, 2007). Bu da bankaların likidite sorunu ile karşı karşıya kalmasına neden olabilmektedir. Bu durumda sorunlu aktiflerin sistem dışına çıkarılması esas yapılması gereken işlemdir. Ancak bankaların zaman zaman sorunlu aktiflerinin yani tahsili gecikmiş alacakların tasviyesi yoluna gitmesi başka sorunları da beraberinde getirmektedir. Bu kredilerin tekrar finanse edilmesi noktasında ise Varlık Yönetim Şirketleri devreye girmektedir. Varlık yönetim şirketleri bankaların zorda kaldığı durumlarda banka aktifinde bulunan sorunlu varlıklarını devralarak bankaları önemli bir sorundan kurtarmaktadır (Tahtakılıç, 2004).

Kredilerin bankacılık sektöründe önemli yeri olduğu tartışılmaz bir gerçektir. Kredilerin sorun teşkil ettiği veya edebileceği riskler dikkate alındığında bankalar kredilendirme işleminin ilk aşamasından itibaren sistemi güvence altına alacak yöntemlerle varlıklarını takip altında tutmaya çalışırlar. Yani kredilerin ilk aşamasından itibaren uzman kişiler tarafından takip edilmesi sağlanmakta ve bu iş bu konuda uzmanlığa sahip olan varlık yönetim şirketleri tarafından yürütülmektedir (Bahadır, 2018). Nitekim tarih boyunca insanların karşılaştıkları ekonomik problemlerin çözümünde farklı dönemlerde farklı mekanizmalar ortaya çıkmıştır. Bu bakımdan modern dönemin ürettiği kurumlar olarak dikkat çekmekte olan Varlık Yönetim Şirketleri alacağın satın alınması, satılması, tahsili, yeniden yapılandırılması ve aracılık hizmetleri vermektedirler (Ak, 2014).

Varlık Yönetim Şirketlerinin geçmişi dünyada 1930'lu yıllara dayanmasına karşın ülkemizde on yıldan biraz fazla bir tarihe sahiptir (Bahadır, 2018, s.102). Ülkemizde 2000 ve 2001 yıllarında yaşanan krizler sonucunda sektörlerde ciddi sıkıntılar oluşmuştur. Bu krizlerin sebep olduğu zararları en aza indirmek ve AB kriterlerinin yerine getirebilmek için kurulan "Üretim ve Finans Danışma Kurulu"nun

getirdiđi iki önemli öneriden birisi “Varlık Yönetim Şirketleri”nin kurulması olarak dikkat çekmektedir (Koç, 2010).

Varlık yönetim şirketleri ilk defa “Mali Sektöre Olan Borçların Yeniden Yapılandırılması ve Bazı Kanunlarda Deđişiklik Yapılması Hakkındaki Kanun” ile sisteme dahil edilmiş ve 5411 sayılı Bankacılık Kanunu’nda yapılan düzenlemeyle devam ettirilmiştir (Tekten, 2018). Günümüzde ise Tasarruf Mevduatı Sigorta Fonu, bankalar ve diđer mali kurumların alacakları ile diđer varlıklarının satın alınması, tahsili, yeniden yapılandırılması ve satılması amacıyla faaliyet göstermek üzere kurulan Varlık yönetim şirketleri Bankacılık Düzenleme ve Denetleme Kurumu (BDDK) tarafından denetlenmekte ve faaliyetlerini sürdürmektedir. 2006 yılında yayımlanan “Varlık Yönetim Şirketlerinin Kuruluş Ve Faaliyet Esasları Hakkında Yönetmelik”te Varlık Yönetim Şirketleri’nin faaliyet alanları;

- Banka, Tasarruf Mevduatı Sigorta Fonu, diđer mali kurumlar ve kredi sigortası hizmeti veren sigorta şirketlerinin söz konusu hizmetlerinden doğan alacakları ile diđer varlıklarını satın alabilir, satabilir, satın aldığı alacakları tahsil edebilir, varlıkları nakde çevirebilir veya bunları yeniden yapılandırarak satabilir.
- Alacakların tahsili amacıyla edindiđi gayrimenkul veya sair mal, hak ve varlıkları işletebilir, kiralayabilir ve bunlara yatırım yapabilir.
- Alacaklarını tahsil etmek amacıyla borçlularına ilave finansman sağlayabilir.
- Banka, Tasarruf Mevduatı Sigorta Fonu, diđer mali kurumlar ve kredi sigortası hizmeti veren sigorta şirketlerinin söz konusu hizmetlerinden doğan alacakları ile diđer varlıklarının tahsilatı, yeniden yapılandırılması veya üçüncü kişilere satışı konularında aracılık, destek ve danışmanlık hizmeti verebilir.
- Ana faaliyetlerini gerçekleştirmek üzere sermaye piyasası mevzuatı dâhilinde ve gerekli izinleri almak kaydıyla faaliyette bulunabilir ve menkul kıymet ihraç edebilir, ihraç edilmiş menkul kıymetlere yatırım yapabilir.
- Faaliyetlerini gerçekleştirmek amacıyla iştirak edebilir.
- Şirketlere kurumsal ve finansal yeniden yapılandırma alanlarında danışmanlık hizmeti verebilir.

şeklinde ifade edilmektedir.

Varlık Yönetim Şirketleri bankaların aktifindeki sorunlu varlıkları rayiç bedeli üzerinden satın alıp daha likit varlıklar olarak geri verirler. Bu işlemi yaparken de alacağı yasal yoluyla takip etmek yerine, yeniden yapılandırma ve borçluyu yeniden finanse etme yöntemlerini kullanarak tahsil etme yoluyla borcun geri ödenmesi sağlamaktadırlar (Tahtakılıç, 2004). Bu araştırmada da varlık yönetim şirketi olarak, borcun geri ödenmesindeki değişkenler incelenerek, varlık yönetimi şirketlerinin ileriye dönük eylem planı oluşturmalarına katkı sağlanması hedeflenmektedir.

İkinci bölümde; önce bu çalışma ile benzerlik gösteren çalışmalar incelenmiştir. Ardından da veri madenciliği, veri madenciliği yöntemleri, sınıflandırma teknikleri, karar ağaçları, Naive Bayes, nitelik seçimi ve varlık yönetimi şirketleri ile ilgili literatüre yer verilmiştir.

3. GEREÇ VE YÖNTEMLER

Çalışmada varlık yönetimi şirketlerine ait borçluların cinsiyet, yaşı, doğum yeri, merkez bankası riski, alacaklı sayısı, sigortasının olması, malvarlığının olması, kefile sahip olması, aile bireylerine ait malvarlığı olması, anapara, toplam bakiye, toplam tahsilat, ikamet ettiği il, ikamet ettiği ilçe, ikamet ettiği mahalle, ülke gelir skoru, il gelir skoru, ilçe gelir skoru, ödeme eğilimi ve profesyonel bir şirketin hazırlamış olduğu adres skorları üzerinden müşteri kategorisi belirlenerek ilgili değişkenlerin hangilerinin ne kadar önemli olduğu incelenmiştir. Bu çalışmayı gerçekleştirebilmek için kullanılacak verilerdeki borçlu seçimi, Türkiye’de varlık yönetim şirketi olarak faaliyet gösteren Gelecek Varlık Yönetimi şirketinin borçlularından seçilmiştir. Varlık yönetim şirketleri devir aldıklarını borçluların bilgilerinin bir kısmını bankadan diğer kısımlarını ise Ulusal Yargı Ağı Bilişim Sistemi (UYAP) ve Merkezî Nüfus İdare Sistem (MERNİS) sistemleri üzerinden toplamaktadır. Veri toplama sürecinde önce varlık yönetim şirketinin veri tabanından borçluların demografik niteliklerin (cinsiyet, maaş, yaş, doğum yeri gibi), merkez bankası riski, kefil, malvarlığı, anapara, yapılan tahsilat, bakiye ve profesyonel şirket tarafından hazırlanan lokasyonlara göre verilmiş skorlarla birlikte alınmıştır. Her iki veri arasındaki ilişkiler analiz edilerek bulgulara ulaşılmıştır. Araştırmanın verilerinin analizinde veri madenciliği yöntemlerinden biri olan karar ağaçları yöntemi kullanılmıştır. Niteliklerin içinden kategorik ve numerik tipinden olmasından dolayı C4.5 algoritmasını tercih edilmiştir. Karar ağaçları oluşturmada WEKA programı yardımıyla C4.5 algoritması uygulanmış, uygulama sonrası elde edilen bulgular yorumlanmıştır.

3.1. Veri

Araştırmada, 49.512 borçluya ait 20 adet cinsiyet, yaş, doğum yeri, merkez bankası riski, alacaklı sayısı, sigortasının olması, malvarlığının olması, kefile sahip

olması, aile bireylerine ait malvarlığı olması, anapara, toplam bakiye, toplam tahsilat, ikamet ettiği il, ikamet ettiği ilçe, ikamet ettiği mahalle, ülke gelir skoru, il gelir skoru, ilçe gelir skoru, ödeme eğilimi nitelikleri ile 1 sınıflandırma verisi (müşteri kategorisi) veri olarak incelenmiştir. Verilere ilişkin bilgiler Şekil 3.1’de verilmiştir.

Şekil 3.1. Veri önizlenimi

Araştırmada kullanılan niteliklerin özellikleri ve değerlendirme sırasında alabileceği değerler Çizelge 3.1’de gösterilmiştir.

Çizelge 3.1. Niteliklerin alabileceği değerler

Özellik	Alabileceği Değerler
Cinsiyet	Kadın, Erkek
Doğum Tarihi	Tarihsel Değer
Alacaklı Sayısı	Numerik Değer
Sigorta	Var, Yok
Malvarlığı	Var, Yok
Kefil	Var, Yok
Ailede Malvarlığı	Var, Yok
Merkez Bankası Riski	Parasal Değer
Anapara	Parasal Değer
Toplam Bakiye	Parasal Değer
İkamet Ettiği İl	Şehirler
İkamet Ettiği İlçe	İlçeler
İkamet Ettiği Mahalle	Mahalleler
Ülke Gelir Skoru	Numerik Değer

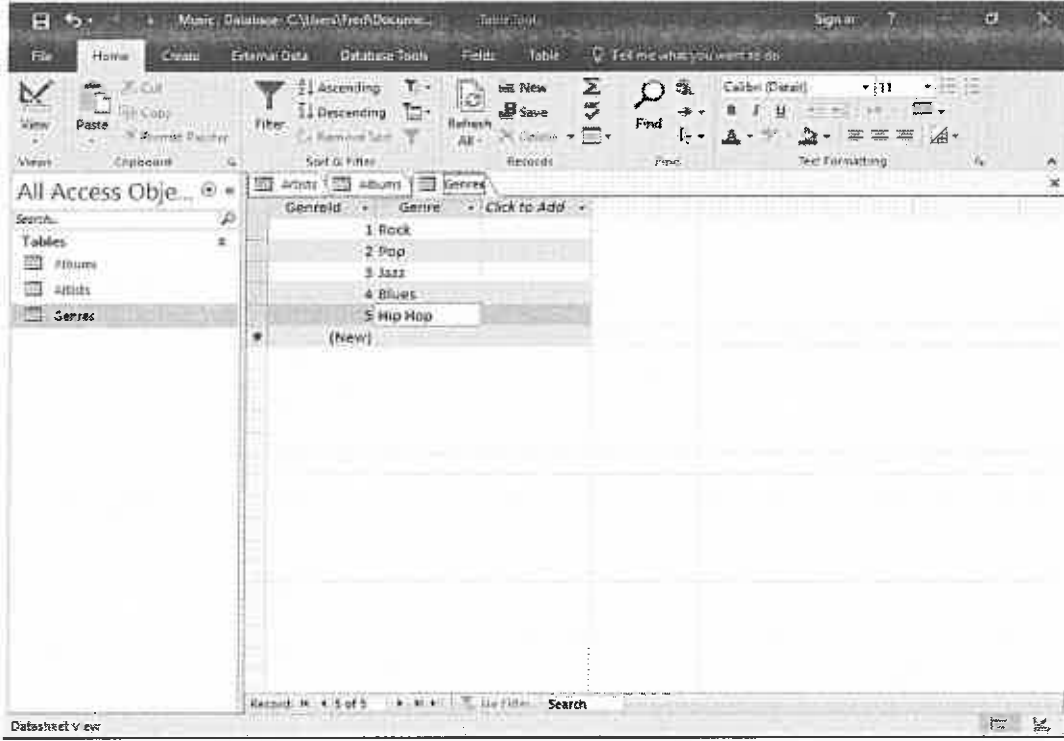
Çizelge 3.2. Niteliklerin alabileceği değerler (devamı)

Özellik	Alabileceği Değerler
İl Gelir Skoru	Numerik Değer
İlçe Gelir Skoru	Numerik Değer
Pasif	Evet, Hayır
Ödeme Eğilimi	odeme gucu yok, odeme niyeti yok, odeme gucu var, odeme niyeti yok, odeme gucu yok, odeme niyeti var, odeme gucu var, odeme niyeti var
Müşteri Kategorisi	tahsilat yapılamayan tahsilat yapılan ödeme planı olan ibra edilen

3.2. Kullanılan Teknolojiler

3.2.1. Microsoft Access

Araştırmada kullanılan verilerin birleştirilmesinde Microsoft Access programı kullanılmıştır. Microsoft Access programı kurulumu ve kullanımı kolay bir program olmasının yanı sıra kolayca entegre edilebilir olması, büyük dosya desteği sunması ve veri ekleme işleminin kolaylıkla yapılması nedeniyle tercih edilmiştir. Microsoft Office'e ait bir program olan Microsoft Access (veya MS Access) bilgisayarda verilerin saklanması ve üzerinde işlem yapılabilmesi için kullanılmaktadır. Microsoft Access bir veritabanı sistemi olup verileri bilgisayara .accdb dosya uzantısı ile kaydetmektedir. Daha önceki Microsoft Access sürümlerinde dosya uzantıları .mdb veya .mde uzantıları kullanılmış olmasına karşın günümüzde Microsoft .accdb dosya uzantısını kullanmaktadır. Araştırmada kullanılan mevcut veriler .accdb dosya uzantısında olduğundan herhangi bir işleme tabi tutulmak zorunda kalmamıştır. Araştırma kullanılan Microsoft Access programının ekran görüntüsü Şekil 3.2'de verilmiştir.



Şekil 3.2. MS ekran görüntüsü

3.2.2. Weka

Araştırmadaki verilerin sınıflandırılması ve değerlendirilmesi için Weka programı kullanılmıştır. Waikato Üniversitesinde geliştirilmiş olan WEKA Programının (Waikato Environment for Knowledge Analysis) 3.6.9 açık kaynak kodlu bir yazılım olup program pek çok sınıflandırma, kümeleme ve birliktelik kurallarına ait algoritmayı desteklemektedir. Ayrıca WEKA, metin tabanlı arff, arff.gz, names, data, csv, c45, libsvm, dat, bsi, xrf, xrf.gz dosya tiplerinin yanı sıra, veri tabanlarını ve verilerin olduğu URL adreslerini de desteklemektedir (Alan, 2014, s. 106).

C4.5, Quinlan tarafından geliştirilmiş ve ID3 algoritmasının bir uzantısı olan bir karar ağacı algoritmasıdır. C4.5 algoritması, numerik ve kategorik verilerle çalışabiliyor olması, hızlı sınıflandırma yapması ve yüksek hassasiyet göstermesi nedeniyle sınıflandırma yapmada yaygın olarak kullanılmaktadır (Sharma, Agrawal ve Sharma, 2013). ID3'teki bir dizi iyileştirme, C4.5 olarak adlandırılan karar ağacı indüksiyonu için pratik ve etkili bir sistemle sonuçlanmıştır. Weka veri madenciliği görevlerinde yönlendiren grafiksel bir kullanıcı arayüzüne ve modellerin anlaşılmasına yardımcı olacak mükemmel veri görselleştirme araçlarına sahiptir. 2011 yılında Weka'nın geliştiricileri makine öğrenme yazılımını C4.5'i "muhtemelen bugüne

kadar pratikte en yaygın kullanılan makine öğrenme aracı programı” olarak ifade etmişlerdir (Witten ve Frank, 2005).

Verilerin hazırlanması, sınıflandırılması, değerlendirilmesi ve regresyonu gibi birçok araca sahip olan Weka bir algoritma topluluğu olup veri madenciliğinde de kullanılabilen Yeni Zelanda'da Waikato Üniversitesi tarafından geliştirilen bir veri madenciliği uygulamasıdır. Weka, GNU (Genel Açık Lisans) lisansı altında yayınlanmış açık kaynaklı bir yazılım olup uygulamanın kurulabilmesi için gerekli olan Java versiyonları Şekil 3.3’de verilmiştir.

		Java				
		1.4	1.5	1.6	1.7	1.8
WEKA	<3.4.0	X	X	X	X	X
	3.4.x	X	X	X	X	X
	3.5.x	3.5.0-3.5.2	>3.5.2	X	X	X
	3.6.x		X	X	X	X
	3.7.x		3.7.0	>3.7.0	>3.7.13	X
	3.8.x				3.8.0-3.8.1	>3.8.1
	3.9.x				3.9.0-3.9.1	>3.9.1

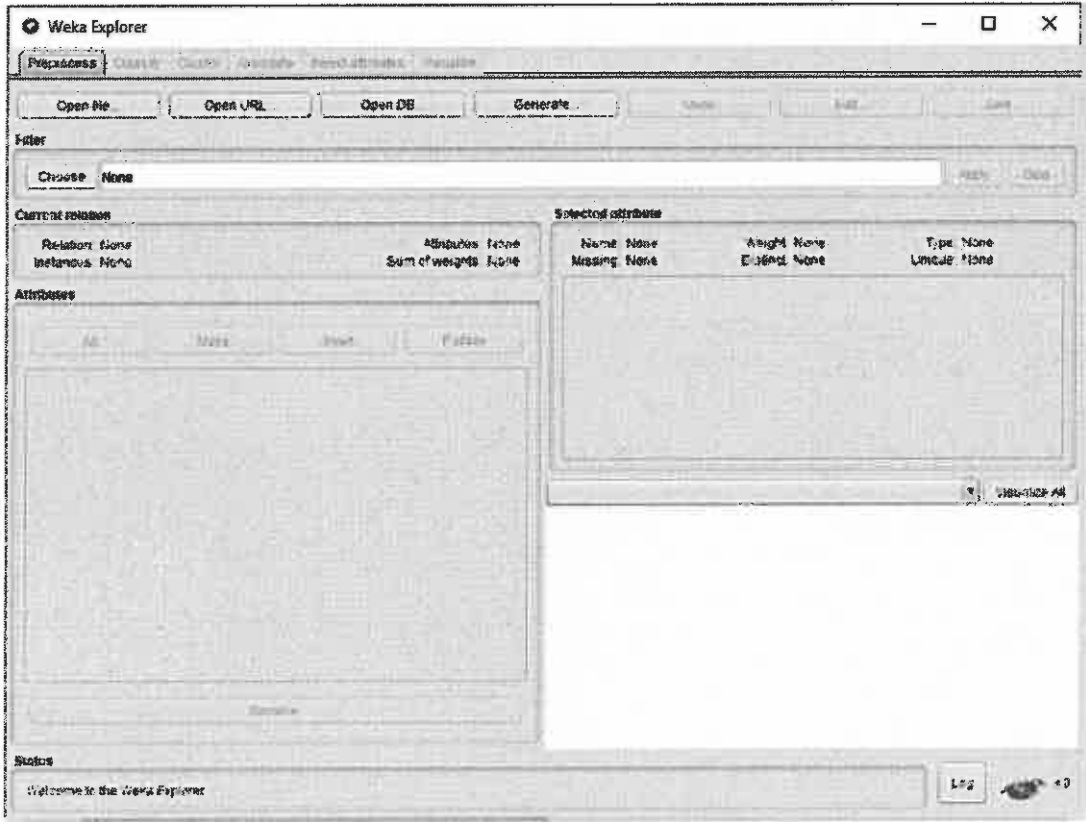
Şekil 3.3. Weka uygulaması java gereksinimi

Weka programının giriş ekranında Keşfet, Araştırmacı, Bilgi Akışı, Çalışma Tezgahı ve Basit Komut Satırı Arayüzü seçenekleri Şekil 3.4’de verilmiştir.



Şekil 3.4. Weka giriş ekranı

Keşfet (Explorer) özelliği verilerin keşfedilebilmesi için bir ortam sağladığından araştırmanın verilerinin sınıflandırma işlemi sırasında Weka programının Keşfet (Explorer) ekranından yararlanılmıştır. Bu ekranda verilerin yüklenmesi, düzenlenmesi, hangi verilerle çalışılacağı, sınıflandırma metodu ve özellik seçimi gibi işlemler yapılmaktadır. Weka Keşfet ekran görüntüsü Şekil 3.5’de verilmiştir.



Şekil 3.5. Weka keşfet ekranı

3.3. Veri Toplama

Araştırmaya ilişkin veriler biri varlık yönetimi alanında biri de yazılım alanında faaliyet gösteren iki firmanın incelemek ve değerlendirmek için oluşturdukları havuzda yer alan bilgilerden alınmıştır. Araştırmada analiz edilen veriler varlık yönetim şirketinin ihale işlemi sonucunda devir aldığı 52.159 borçluya ait bilgiler ile profesyonel bir yazılım şirketinin sağlamış olduğu 521.159 adres skorundan oluşmaktadır. Veriler birleştirilip borçlular üzerinde sadece bireysel borçlulara yönelik filtreleme yapıldığında kalan 49.512 borçluya ait veri elde edilmiştir.

Arařtırmada yer alan borçlulara ait bilgilerine iliřkin ‘Debtor’ Çizelgesunda; cinsiyet, yař, doęum yeri, merkez bankası riski, alacaklı sayısı, sigortasının olması, malvarlığının olması, kefile sahip olması, aile bireyelerine ait malvarlığı olması, anapara, toplam bakiye, toplam tahsilat alanları yer almaktadır. Varlık yönetimi řirketinden alınan borçlulara ait bilgilerin yer aldığı borçlu Çizelgesunun veri türleri Şekil 3.6’de verilmiştir.

Field Name	Data Type
SEX	Number
BIRTH_DATE	Date/Time
BIRTH_PLACE	Short Text
CB_RISK_LOC	Number
CREDITOR_COUNT	Number
SGK_FLAG	Yes/No
ASSET_OWNER_FLAG	Yes/No
GUARANTOR_HAS_ASSET	Yes/No
FAMILY_HAS_ASSET	Yes/No
CUST_CATEGORY	Number
CAT_DESC	Short Text
POTENTIAL_LEVEL	Number
PO_DESC	Short Text
MAIN_RISK_LOC	Number
TOTAL_RISK_LOC	Number
TOTAL_COLL_LOC	Number
INACTIVE_FLAG	Yes/No
ADDRESS_ID	Short Text

Şekil 3.6. Borçlu Çizelgesu

Microsoft Access programında tutulan borçlu Çizelgesunda yer alan verilerin önizlemesi Şekil 3.7’de verilmiştir.

SEX	BIRTH_DATE	BIRTH_PLACE	CR_RISK_SCORE	CREDITOR_C	SCORE_FLAG	ASSET_OWN	GUARANTOR	FAMILY_HAZ
0	0.03.1986	YILDIZLI	7620	4	0	0	0	0
0	15.04.1973	ŞEBİNKARAHİSAR	12250	3	0	0	0	0
0	10.09.1967	BAFRA	6339	7	0	0	0	0
0	16.12.1979	ERURU	76316	8	0	0	0	0
0	2.03.1974	KAYSERİ	2624	2	0	0	0	0
0	28.03.1962	MALATYA	5156	2	0	0	0	0
0	2.01.1978	GÜRENE	4638	2	0	0	0	0
0	24.01.1975	AYDIN	8457	2	0	0	0	0
0	24.02.1971	KAYSERİ	9940	2	0	0	0	0
0	4.10.1972	İSTANBUL	17760	5	0	0	0	0
0	25.03.1971	ÇİÇEKDAĞI	2369	3	0	0	0	0
0	3.09.1969	İSTANBUL	1373	1	0	0	0	0
0	31.07.1978	AŞKALE	503139	7	0	0	0	0
0	1.01.1964	ARPAÇAY	10121	5	1	0	0	0
0	20.04.1972	MAÇKA	6137	3	0	0	0	0
0	20.12.1978	GAZİANTEP	6381	3	0	0	0	0
0	17.11.1974	ZONGULDAK	20419	6	0	0	0	0
0	7.04.1966	AMASYA	30700	3	0	0	0	0
0	2.06.1961	ZONGULDAK	3739	4	0	0	0	0
0	1.01.1973	HİNİS	68511	9	0	0	0	0
0	16.10.1974	KARABÜK	23239	3	0	0	0	0

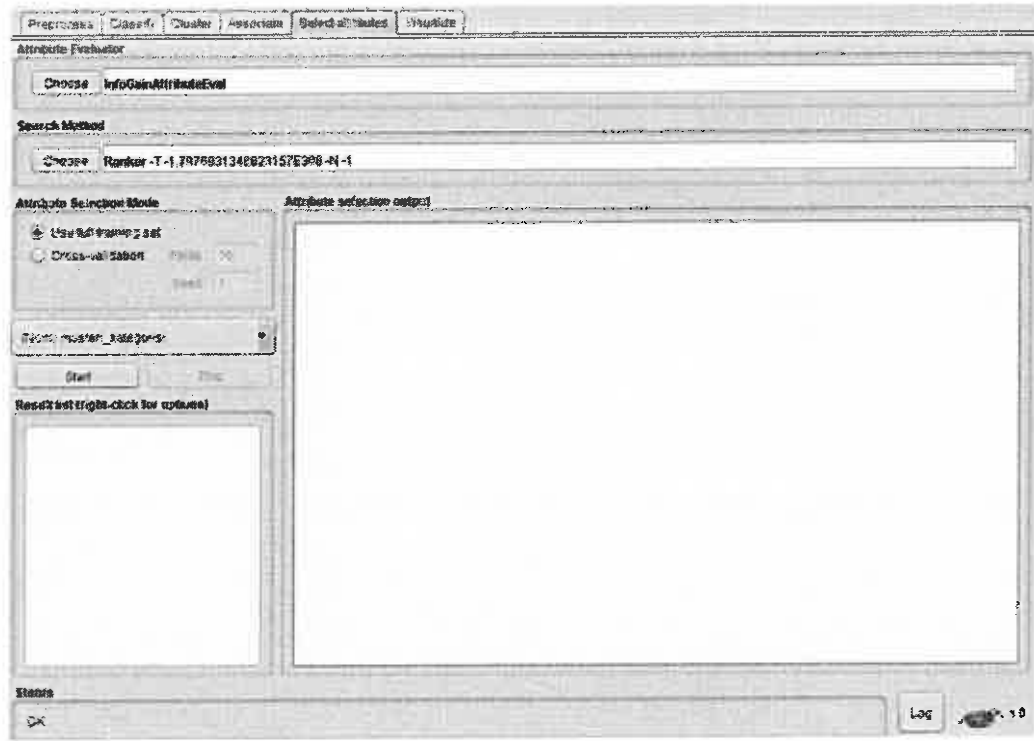
Şekil 3.7. Borçlu Çizelgesi önizleme

Yazılım şirketinden alınan ve borçluların adres skorlarının yer aldığı 'Address' Çizelgesinde lokasyon, ülke gelir skoru, il gelir skoru, ilçe gelir skoru alanları yer almaktadır. Borçluların adres skorlarına ilişkin bilgiler Şekil 3.8'de verilmiştir.

Field Name	Data Type
address_id	Short Text
reidin_il	Short Text
reidin_ilce	Short Text
reidin_mah	Short Text
latitide	Number
longitude	Number
TR_Gelir_Score	Number
TR_Gelir_Guven	Number
Il_Gelir_Score	Number
Il_Gelir_Guven	Number
Ilce_Gelir_Score	Number
Ilce_Gelir_Guven	Number
gelir	Number
kiram2	Number
satıs	Number
SED_Siralama	Number
SED	Short Text

Şekil 3.8. Adres Çizelgesi

Microsoft Access programında tutulan adres Çizelgesinde yer alan verilerin önizlemesi Şekil 3.9'de verilmiştir.



Şekil 3.10. Weka ile nitelik seçme ekranı

Araştırmanın analizine başlamadan önce nitelikler (cinsiyet, yaş, doğum yeri, merkez bankası riski, alacaklı sayısı, anapara, toplam bakiye, toplam tahsilat, ikamet il, ikamet ilçe, ülke gelir skoru, il gelir skoru) ile Bilgi Kazanç Nitelikleri (Information Gain Attributes) algoritmasıyla beraber incelenmiştir. Borçluların niteliklerinden ülke gelir skoru, il gelir skoru, ilçe gelir skoru, kefil, malvarlığı, ailedeki mal varlığının sonucu çok etkilemediği öte yandan toplam tahsilat, toplam bakiye, ikamet ettiği mahalle, doğum yeri ise karar ağacını ciddi oranda etkilediği görülmüştür. Ayrıca lokasyon skor bilgilerinin sonucu etkilemediği ancak ikamet ettiği il, ikamet ettiği ilçe ve ikamet ettiği mahalle bilgilerinin sonucu etkilediği tespit edilmiştir. Karar ağacının oluşmasına etki eden nitelikler Şekil 12’de verilmiştir.

Ranked attributes:

0.8366924	12	toplam_tahsilat
0.2230563	11	toplam_bakiye
0.1848463	15	ikamet_mahalle
0.1841073	3	dogum_yeri
0.0625039	14	ikamet_ilce
0.0521595	19	pasif
0.0407041	4	merkez_bankasi_riski
0.0389064	20	odeme_egilimi
0.0380729	10	anapara
0.0249979	5	alacakli_sayisi
0.0132495	13	ikamet_il
0.0115498	2	yas
0.0088596	6	sigorta
0.0014336	1	cinsiyet
0.0002779	9	ailede_malvarligi
0.0002461	7	malvarligi
0.0000455	8	kefil
0	18	ilce_gelir_skoru
0	16	ulke_gelir_skoru
0	17	il_gelir_skoru

Selected attributes: 12,11,15,3,14,19,4,20,10,5,13,2,6,1,9,7,6,18,16,17 : 20

Şekil 3.11. Niteliklerin sıralanma sonucu

Araştırmada kullanılan cinsiyet, yaş, doğum yeri, alacaklı sayısı, anapara, merkez bankası riski, ödeme eğilimi, sigorta, toplam bakiye, toplam tahsilat, ikamet ettiği il, ikamet ettiği ilçe, ikamet ettiği mahalle gibi niteliklerin veri dağılımları Çizelgelendirilmiştir.

Araştırmanın verilerinde 34320 kişinin tahsilat yapılamayan 34320 borçlu, 11023 kişinin tahsilat yapılan borçlu, 3876 kişinin ödeme planı olan borçlu ve 292 kişinin ibra edilen borçlu olduğu görülmüştür. Araştırmada kullanılan verilerin, müşteri kategorisine göre dağılımı Çizelge 3.2’de yer verilmiştir.

Çizelge 3.3. Müşteri kategorisi (sınıf) veri dağılımı

No.	Kategoriler	Sayı
1.	Tahsilat Yapılamayan	34320
2.	Tahsilat Yapılan	11023
3.	Ödeme Planı Olan	3876
4.	İbra Edilen	292

Arařtırmada 43679 kadın borçlu ve 5832 kiřinin erkek borçluya ait veri bulunmaktadır. Arařtırmada kullanılan verilerin, cinsiyet deęiřkenine göre daęılımı Çizelge 3.3’de verilmiřtir.

Çizelge 3.4. Borçluların cinsiyet daęılımı

No.	Cinsiyet	Sayı
1.	Kadın	43679
2.	Erkek	5832

Arařtırmanın verilerinde borçluların yařlarının minimum 21, maksimum 119, ortalamasının 46,086 ve standart sapmasının 10,021 olduęu görölmüřtür. Arařtırmada kullanılan verilerin, yař deęiřkenine göre daęılımı Çizelge 3.4’de verilmiřtir.

Çizelge 3.5. Borçluların yař daęılımı

No.	İstatistiksel Ölçütler	Deęer
1.	Minimum	21
2.	Maksimum	119
3.	Ortalama	46,086
4.	Standart Sapma	10,021

Arařtırmadaki borçluların doęum yerleri gruplandırılmıřtır. Gruplama iřlemi aynı yerde bulunan borçlu sayısına göre yapılmıřtır. Borçluların doęum yeri bilgisi aısından 1-500 arasında borçlu bulunan 3529 doęum yeri, 501-1000 arasında borçlu bulunan 3 doęum yeri, 1001-1500 arasında borçlu bulunan 1 doęum yeri, 1501-2000 arasında borçlu bulunan 1 doęum yeri, 2001-2500 arasında borçlu bulunan 0 doęum yeri, 2501-300 arasında borçlu bulunan ise 1 doęum yeri olduęu görölmüřtür.

Arařtırmada kullanılan verilerin, doęum yeri deęiřkenine gre daęılımını izelge 3.5’de verilmiřtir.

izelge 3.6. Borluların doęum yeri daęılımını

No.	Daęılım Aralıkları	Deęer
1.	1-500	3529
2.	501-1000	3
3.	1001-1500	1
4.	1501-2000	1
5.	2001-2500	0
6.	2501-3000	1

Arařtırmada 0 kiři alacaklı olan 3, 1 kiři alacaklısı olan 7570, 2 tane alacaklısı olan 11702, 3 tane alacaklısı olan 11304, 4 tane alacaklısı olan 8600, 5 tane alacaklısı olan 5571, 6 tane alacaklısı olan 3603, 7 tane alacaklısı olan 1975, 8 tane tane alacaklısı olan 1077, 9 tane alacaklısı olan 447, 10 tane alacaklısı olan 178, 11 tane alacaklısı olan 94, 12 tane alacaklısı olan 13, 13 tane alacaklısı olan 18 ve 14 tane alacaklısı olan 4 borlu olduęu tespit edilmiřtir. Arařtırmada kullanılan verilerin, alacaklı sayısı deęiřkenine gre daęılımını izelge 3.6’de verilmiřtir.

izelge 3.7. Borluların alacaklı sayısı daęılımını

No.	Alacaklı sayısı	Toplam
1.	0	3
2.	1	7570
3.	2	11702
4.	3	11304
5.	4	8600

Çizelge 3.8. Borçluların alacaklı sayısı dağılımı (devamı)

No.	Alacaklı sayısı	Toplam
6.	5	5571
7.	6	3603
8.	7	1975
9.	8	1077
10.	9	447
11.	10	178
12.	11	94
13.	12	13
14.	13	18
15.	14	4

Araştırmada borçlu anaparasının minimum 0 TL, maksimum 250.377,59 TL, ortalama 12508,885 TL, standart sapma 116172,149 olduğu görülmüştür. Araştırmada kullanılan verilerin, anapara değişkenine göre dağılımı Çizelge 3.7’de verilmiştir.

Çizelge 3.9. Borçluların anapara dağılımı

No.	İstatistikse ölçütler	Değer (TL)
1.	Minimum	0
2.	Maximum	250.377,59
3.	Ortalama	12.508,885
4.	Standart Sapma	116172,149

Borçlu anaparasının minimum 11 TL, maksimum 2.011.157 TL, ortalama 35.653,671 TL ve standart sapmanın 56945,209 olduğu görülmüştür. Araştırmada kullanılan verilerin, merkez bankası riski değişkenine göre dağılımı Çizelge 3.8’de verilmiştir.

Çizelge 3.10. Borçluların merkez bankası riski dağılımı

No.	İstatistiksel Ölçütler	Değer (TL)
1.	Minimum	11
2.	Maximum	2.011.157
3.	Ortalama	35.653,671
4.	Standart Sapma	56.945,209

Araştırmada “ödeme gücü yok ve ödeme niyeti yok” olan 1685 borçlu, “ödeme gücü yok ve ödeme niyeti var” olan 5832 borçlu, “ödeme gücü var ve ödeme niyeti yok” olan 195572 borçlu, “ödeme gücü var ve ödeme niyeti var” olan 12335 borçlu olduğu görülmüştür. Araştırmada kullanılan verilerin, ödeme eğilimi değişkenine göre dağılımı Çizelge 3.9’de verilmiştir.

Çizelge 3.11. Borçluların ödeme eğilimi dağılımı

No.	Ödeme Eğilimi	Sayı
1.	Ödeme gücü yok, Ödeme niyeti yok	1685
2.	Ödeme gücü yok, Ödeme niyeti var	5832
3.	Ödeme gücü var, Ödeme niyeti yok	195572
4.	Ödeme gücü var, Ödeme niyeti var	12335

Araştırmada sigortası olmayan 46554 borçlu, sigortası olan 2957 borçlu olduğu görülmüştür. Araştırmada kullanılan verilerin, sigorta değişkenine göre dağılımı Çizelge 3.10’de verilmiştir.

Çizelge 3.12. Borçluların sigorta dağılımı

No.	Sigorta Bilgisi	Sayı
1.	Hayır	46554
2.	Evet	2957

Araştırmada borçlu toplam bakiyesinin minimum 11 TL, maksimum 2.011.157 TL, ortalama 35.653,671 TL, standart sapma 56945,209 olduğu görülmüştür. Araştırmada kullanılan verilerin, toplam bakiye değişkenine göre dağılımı Çizelge 3.11’de verilmiştir.

Çizelge 3.13. Borçluların toplam bakiye dağılımı

No.	Bakiye İstatistikleri	Değer (TL)
1.	Minimum	11
2.	Maximum	2.011.157
3.	Ortalama	35.653,671
4.	Standart Sapma	56.945,209

Araştırmada borçlu toplam tahsilatının minimum -1000.05 TL, maksimum 155.000 TL, ortalama 1083,882 TL, standart sapma 3980,067 olduğu görülmüştür. Araştırmada kullanılan verilerin, toplam tahsilat değişkenine göre dağılımı Çizelge 3.12’de verilmiştir.

Çizelge 3.14. Borçluların toplam tahsilat dağılımı

No.	İstatistiksel Ölçütler	Değer (TL)
1.	Minimum	-1000.05
2.	Maximum	155.000
3.	Ortalama	1.083,882
4.	Standart Sapma	3.980, 067

Araştırmada borçluların ikamet ettiği illeri gruplanarak gösterilmiştir. Gruplama işlemi verideki bulunan borçlu sayısına göre yapılmıştır. 1-500 borçlu bulunan 60 ikamet edilen il, 501-1000 arasında bulunan 12 ikamet edilen il, 1001-1500 arasında bulunan 3 ikamet edilen il, 1501-2000 arasında bulunan 2 ikamet edilen il, 2001-2500 arasında bulunan 2 ikamet edilen il, 2501 ve fazlasında 3 tane ikamet edilen il olduğu görülmüştür. Araştırmada kullanılan verilerin, ikamet ettiği il değişkenine göre dağılımı Çizelge 3.13’de verilmiştir.

Çizelge 3.15. Borçluların ikamet ettiği il dağılımı

No.	Dağılım Aralıkları	Değer
1.	1-500	60
2.	501-1000	12
3.	1001-1500	3
4.	1501-2000	2
5.	2001-2500	2
6.	2501-	3

Araştırmada borçluların ikamet ettiği ilçeler gruplanarak gösterilmiştir. Gruplama işlemi verideki bulunan borçlu sayısına göre yapılmıştır. 1-500 borçlu bulunan 855 ikamet edilen ilçe, 501-1000 arasında bulunan 16 ikamet edilen ilçe, 1001-1500 arasında bulunan 1 ikamet edilen ilçe, 1501-2000 arasında bulunan 0 ikamet edilen ilçe, 2001-2500 arasında bulunan 0 ikamet edilen ilçe, 2501-3000

arasında bulunan 0 ikamet edilen ilçe, 3000 ve fazlasında ise 1 ikamet edilen ilçe olduğu görülmüştür. Araştırmada kullanılan verilerin, ikamet ettiği ilçe değişkenine göre dağılımı Çizelge 3.14’de verilmiştir.

Çizelge 3.16. Borçluların ikamet ettiği ilçe veri dağılımı

No.	Dağılım Aralıkları	Değer
1.	1-500	855
2.	501-1000	16
3.	1001-1500	1
4.	1501-2000	0
5.	2001-2500	0
6.	2501-3000	0
7.	3000-	1

Araştırmada borçluların ikamet ettiği mahalleler gruplanarak gösterilmiştir. Gruplama işlemi veride bulunan borçlu sayısına göre yapılmıştır. 1-500 borçlu bulunan 5716 ikamet edilen mahalle, 501-1000 arasında bulunan 3 ikamet edilen mahalle, 1001-1500 arasında bulunan 0 ikamet edilen mahalle, 1501-2000 arasında bulunan 0 ikamet edilen mahalle, 2001-2500 arasında bulunan 0 ikamet edilen mahalle, 2501-3000 arasında 0 ikamet edilen mahalle, 3000 ve fazlasında ise 1 ikamet edilen mahalle olduğu görülmüştür. Araştırmada kullanılan verilerin, ikamet ettiği mahalle değişkenine göre dağılımı Çizelge 3.15’de verilmiştir.

Çizelge 3.17. Borçluların ikamet ettiği mahalle dağılımı

No.	Dağılım Aralıkları	Değer
1.	1-500	5716
2.	501-1000	3
3.	1001-1500	0

Çizelge 3.18. Borçluların ikamet ettiği mahalle dağılımı (devamı)

No.	Dağılım Aralıkları	Değer
4.	1501-2000	0
5.	2001-2500	0
6.	2501-3000	0
7.	3000-	1

Üçüncü bölümde; önce araştırmanın amacına yönelik olarak elde edilen veriler ve kullanılan teknolojiler (Microsoft Access, Weka) ardından da veri toplama ve analiz süreci hakkında bilgi verilmiştir.

4. BULGULAR

4.1. C4.5

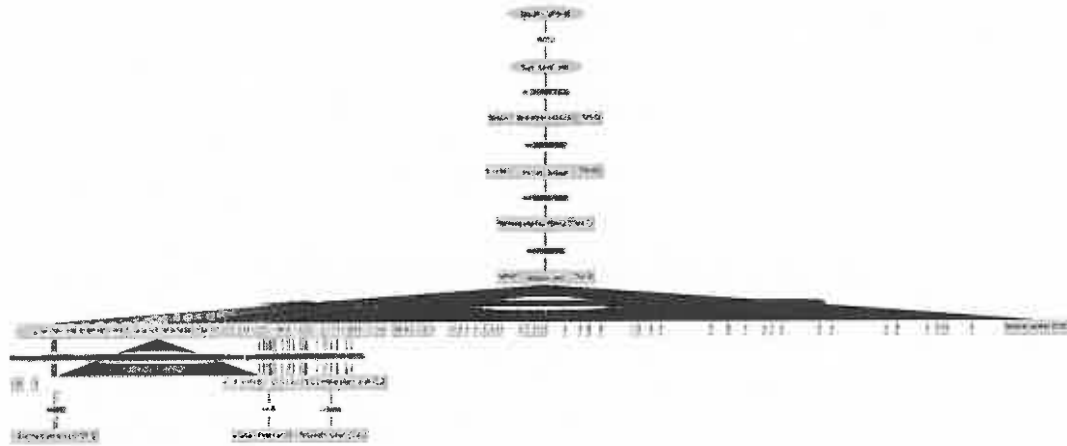
Araştırmanın amaçlarına yönelik olarak verilerin analizini C4.5 karar ağacını kullanılarak elde edilen bulgular bu bölümde yer almaktadır.

Araştırmanın verilerini oluşturan borçluların, borçlarını geri ödeme eğilimlerini inceleyip karar ağacını oluşturabilmek amacıyla cinsiyet, yaş, doğum yeri, merkez bankası riski, alacaklı sayısı, anapara, toplam bakiye, toplam tahsilat, ikamet il, ikamet ilçe, ülke gelir skoru, il gelir skoru gibi nitelikleri C4.5 (Weka'da J48) algoritmasında 10 kez çapraz doğrulama işlemi uygulanmıştır. C4.5 algoritmasıyla oluşan karar ağacında 4275 adet dal ve ağacın boyutunun ise 4323 olarak bulunmuştur. Toplam 49511 tane veriden 45856 (%92.6178)'inin başarılı bir şekilde sınıflandırılmış olduğu 3655 (%7.3822)'inin ise doğru sınıflandırılmadığı görülmektedir. Ayrıca kök ortalama kare hatası 0.1654, görelî mutlak hatası %23.1893, kök görelî kare hatası %47.568 olduğu tespit edilmiştir. C4.5 Karar ağacı sınıflandırma işlemi sonucunda verilerin sınıflandırılmasına yönelik özet bilgiler Çizelge 4.1'de verilmiştir.

Çizelge 4.1. C4.5 Sınıflandırma Sonucu Özeti

Sınıflandırma İstatikleri	Değer	Oran (%)
Başarılı Sınıflandırma Örnekleme	45856	92.7107
Hatalı Sınıflandırma Örnekleme	3609	7.2893
Kappa İstatistiği	0.8384	
Ortalama mutlak hata:	0.0527	
Kök ortalama kare hatası	0.1639	
Görelî mutlak hatası	22.7048 %	
Kök görelî kare hatası	48.1404 %	
Toplam örnek sayısı	49511	

C4.5 Karar ağacı sınıflandırma işlemi sonucunda elde edilen karar ağacı Şekil 4.1'da verilmiştir.



Şekil 4.1. Karar ağacı

Araştırmanın verilerinin analizi sonucunda oluşturulan karar ağacındaki kök niteliğın toplam tahsilat olduđu görülmektedir. Verilerde doğum yeri ve lokasyon bilgilerinin fazla olmasından dolayı karar ağacı 4.11'daki gibi gösterilmiş, karar

ağacının detaylı gösterimi CD'deki EK-1'de verilmiştir. Karar ağacı incelendiğınerek aşağıda verilen 51 adet Eğer-İse biçiminde kurallar oluşturulmuştur:

1. Eğer “Toplam Tahsilat ≤ 10 TL” ve “Toplam Bakiye $\leq 24800,003$ TL” ve “Anapara $\leq 12235,35$ TL” ise “Müşteri Kategorisi = Tahsilat Yapılamayan”.
2. Eğer “Toplam Tahsilat ≤ 10 TL” ve “Toplam Bakiye $\leq 24800,003$ TL” ve “Anapara $> 12235,35$ TL” ve “Toplam Bakiye $\leq 16705,7193$ TL” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
3. Eğer “Toplam Tahsilat ≤ 10 TL” ve “Toplam Bakiye $\leq 24800,003$ TL” ve “Anapara $> 12235,35$ TL” ve “Toplam Bakiye $> 16705,7193$ TL” ise “Müşteri Kategorisi = Tahsilat Yapılamayan”.
4. Eğer “Toplam Tahsilat ≤ 10 TL” ve “Toplam Bakiye $\leq 24800,003$ TL” ve “Anapara $> 12235,35$ TL” ve “Toplam Bakiye $> 16705,7193$ TL” ve “Anapara $> 17182,46$ TL” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
5. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 24800,003$ TL” ise “Müşteri Kategorisi = Tahsilat Yapılamayan”.
6. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $\leq 2529,9997$ TL” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
7. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Anapara $\leq 4839,79$ TL” ise “Müşteri Kategorisi=Tahsilat Yapılan”.
8. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = İstanbul” ve “Anapara $\leq 9335,48$ TL” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
9. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = İstanbul” ve “Anapara $> 9335,48$ TL” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
10. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Ağrı” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
11. Eğer “Toplam Tahsilat >10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Ladik” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
12. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve

- “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Sarıyer” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
13. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Adana” ve “Anapara $\leq 8667,5$ TL” ve “Yaş ≤ 37 ” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
14. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Adana” ve “Anapara $\leq 8667,5$ TL” ve “Yaş > 37 ” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
15. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Adana” ve “Anapara $> 8667,5$ TL” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
16. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Üsküdar” ve “Anapara $\leq 13224,88$ TL” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
17. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Bakırköy” ve “Alacaklı Sayısı ≤ 2 ” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
18. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Bakırköy” ve “Alacaklı Sayısı > 2 ” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
19. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Diyarbakır” ise “Müşteri Kategorisi = Tahsilat Yapılan” (Borçlunun ikamet ettiği ilçe alt kırılımı olmasına rağmen).
20. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Kadıköy” ve “Yaş ≤ 34 ” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
21. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye $> 2529,9997$ TL” ve “Toplam Bakiye $\leq 16801,964$ TL” ve “Anapara $> 4839,79$ TL” ve “Doğum Yeri = Kadıköy” ve “Yaş > 34 ” ise “Müşteri Kategorisi = Ödeme Planı Olan”.

22. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Şanhurfa” ve “Yaş <= 46” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
23. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Şanhurfa” ve “Yaş > 46” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
24. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Araban” ve “Cinsiyet = Kadın” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
25. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Araban” ve “Cinsiyet = Erkek” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
26. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Ankara” ve “Merkez Bankası Riski <= 9315 TL” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
27. Eğer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Ankara” ve “Merkez Bankası Riski <= 9315 TL” ve “Alacaklı Sayısı <= 4” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
28. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Ankara” ve “Merkez Bankası Riski > 9315 TL” ve “Alacaklı Sayısı > 4” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
29. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Kütahya” ve “Yaş <= 43” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
30. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Kütahya” ve “Yaş > 43” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
31. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve

- “Doğum Yeri = Kahramanmaraş” ve “Yaş \leq 41” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
32. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum Yeri = Kahramanmaraş” ve “Yaş $>$ 41” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
33. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum Yeri = Manisa” ve “Merkez Bankası \leq 17145 TL” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
34. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum Yeri = Manisa” ve “Merkez Bankası Riski $>$ 17145 TL” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
35. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum yeri = Gaziantep” ve “Yaş \leq 45” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
36. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum Yeri = Gaziantep” ve “Yaş $>$ 41” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
37. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum Yeri = Antalya” ve “Alacaklı Sayısı \leq 1” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
38. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum Yeri = Antalya” ve “Alacaklı Sayısı $>$ 1” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
39. Eğer “Toplam Tahsilat $>$ 10 TL’den büyük” ve “Toplam Bakiye $>$ 2529,9997 TL” ve “Toplam Bakiye \leq 16801,964 TL” ve “Anapara $>$ 4839,79 TL” ve “Doğum Yeri = Kadirli” ve “Yaş \leq 52” ise “Müşteri Kategorisi = Tahsilat Yapılan”.

40. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Kadirli” ve “Yaş > 52” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
41. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Eskişehir” ve “Yaş ≤ 42” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
42. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Eskişehir” ve “Yaş ≤ 42” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
43. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Beykoz” ve “Merkez Bankası Riski ≤ 12570 TL” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
44. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Beykoz” ve “Merkez Bankası Riski ≤ 12570 TL” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
45. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = Beykoz” ve “Merkez Bankası Riski > 12570 TL” ise “Müşteri Kategorisi = Ödeme Planı Olan”.
46. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye > 16801,964 TL” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
47. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = İzmit” ve “Yaş ≤ 43” ise “Müşteri Kategorisi = Tahsilat Yapılan”.
48. Eğer “Toplam Tahsilat > 10 TL’den büyük” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye ≤ 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doğum Yeri = İzmit” ve “Yaş > 43” ise “Müşteri Kategorisi = Ödeme Planı Olan”.

49. Eđer “Toplam Tahsilat > 10 TL’den byk” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doęum Yeri = Kırıkkale” ve “Merkez Bankası Riski <= 8165 TL” ise “Mşteri Kategorisi = Tahsilat Yapılan”.
50. Eđer “Toplam Tahsilat > 10 TL’den byk” ve “Toplam Bakiye > 2529,9997 TL” ve “Toplam Bakiye <= 16801,964 TL” ve “Anapara > 4839,79 TL” ve “Doęum Yeri = Kırıkkale” ve “Merkez Bankası Riski > 8165 TL” ise “Mşteri Kategorisi = deme Plan Olan”.
51. Eđer “Toplam Tahsilat > 10 TL” ve “Toplam Bakiye <= 2529,9997 TL” ve “Toplam Bakiye <= 4,9982 TL” ve “Toplam Bakiye > -4,2297 TL” ise “Mşteri Kategorisi = İbra Edilen”.

Elde edilen kuralların kendi aralarındaki iliřkileri incelendięinde ařaęıdaki yorumlar yapılabilir:

- 2., 3. ve 4. kurallarında grldę gibi anaparası 12235,35 TL’den byk olup toplam bakiyesi 16705,7193 TL’den kk eřit olan borluların deme planı olduęu grlrken, toplam bakiyesi 16705,7193 TL’den byk olan borluların ise anapara alt kırılımına sahip olduęu ve anaparası 17182,46 TL’den kk eřit olanların tahsilat yapılamayan ve anaparası 17182,46 TL’den byk olanlar ise deme planı olan borlular olduęu belirlenmiřtir.
- 13. 14. ve 16. kurallarında grldę zere doęum yeri Adana olan bir borlunun anaparası 4839,79 TL’den byk ve yaşı 37’den kk ise deme planına sahip olduęu, aynı zamanda yaşı 37’den byk olanların ise geri demeyi yaptığı grlmřtir. te yandan řekilde doęum yeri řkdar olan bir borlunun yař iin bir ayırım olmadığı tespit edilmiřtir. Ancak anapara iin bir alt kırılım daha oluřtuęu gzlemlenmiř ve anaparası 13224,88 TL’den kk veya eřit olanların tahsilat yapılan olan iken anaparası 13224,88 TL’den byk olanın ise deme planı olan bir borlu olduęu belirlenmiřtir. Dięer taraftan 19. maddede de grldę zere doęum yeri Diyarbakır olan borlunun ikamet ettięi ileye gre bir alt kırılım olmasına karřın Diyarbakırlı borlularının tamamının geri deme yaptığı belirlenmiřtir.
- Doęum yeri Bakırky olan borluların alacaklı sayısının bir nem oluřturduęu ve alacaklı sayısı 2’den byk veya kk eřit olması durumlarında 17. maddede grldę zere deme planı olan veya geri deme yapan olarak iřaretlendięi tespit edilmiřtir. 20. ve 21. kurallarında grldę zere ise doęum yeri Kadıky olan bir

borçlunun, doğum yeri Adana'da olduğu gibi yaşın bir önem oluşturduğu ve Kadıköylü olan bir borçlunun 34 yaşında veya küçük olması durumunda tahsilat yapılan ama 34 yaşından büyük olması durumunda ise ödeme planı olan bir borçlu olarak işaretlendiği belirlenmiştir.

- İlk bakıldığında yaşı yüksek olan borçluların geri ödemesini yapmış olduğu görülse de 13., 20. ve 22. kuralları detaylı incelendiğinde yaş bilgisinin doğum yeri bilgisiyle bir bütün olduğu, doğum yerine ve yaşa birlikte bakılarak borçluların farklı kategorilerde olduğu tespit edilmiştir. Doğum yeri Adana olan borçlularda 37 yaşından büyükler geri ödeme yaparken, Kadıköy'de doğan ve 34 yaşından küçük olan ile Şanlıurfa'da doğan ve 46 yaşından küçük olan borçluların borçlarını geri ödediği belirlenmiştir.
- 23. ve 24. kurallarına bakıldığında; doğum yeri Araban olan borçlularda cinsiyetin önemli bir nitelik olarak ortaya çıktığı görülmüştür. Araban'da doğan borçluların cinsiyeti kadın olanların tahsilat yapılan, cinsiyeti erkek olanların ise ödeme planı olan olarak kategorilendirildiği tespit edilmiştir.
- Merkez bankası riski, alacaklı sayısı, sigorta ve anapara bilgisilerinin önemli bir nitelik olarak ortaya çıktığı görülmüştür ve buna bağlı olarak borçluların ödeme durumunun değiştiği tespit edilmiştir.

Verilerin sınıflandırma işleminin sonucunda f_1 skor değerlerinin ibra edilen borçlular için 1.00, tahsilat yapılamayan borçlular için 0.0992 ve tahsilat yapılan borçlular için 0.873 olduğu gözlemlenmiştir ama ödeme planı olan borçluları ise verimli sınıflandırma yapamadığı görülmüştür. Doğru pozitif oranlarına (True-Positive Rate) bakıldığında tahsilat yapılamayan bütün borçluları 1.00 oranıyla tahsilat yapılamayan olarak bulduğunu, tahsilat yapılan borçluları 0.951 oranıyla tahsilat yapılan olarak bulduğunu, ödeme planı borçluları 0.209 oranıyla ödeme planı olan bulduğunu, ibra edilen bütün borçluları 1.00 oranıyla ibra edilen olarak bulduğunu görülmüştür. Yanlış pozitif oranlarına (False-Positive Rate) bakıldığında tahsilat yapılamayan olmayan borçluları 0.037 oranıyla tahsilat yapılamayan olarak bulduğunu, tahsilat yapılan olmayan borçluları 0.065 oranıyla tahsilat yapılan olarak bulduğunu, ödeme planı olmayan borçluları 0.012 oranıyla ödeme planı olan bulduğunu, ibra edilen bütün borçluları 0.00 oranıyla ibra edilen olarak bulduğunu görülmüştür. Kesinlik alanına bakıldığında borçluların gerçekten ne kadar miktarı belirttiğimiz kategori olduğunu inceleyecek olursak, tahsilat yapılamayan olarak

belirtilen borçuların 0.984'nün tahsilat yapılamayan olduğunu, tahsilat yapılan olarak belirtilen borçuların 0.807'si tahsilat yapılan olduğunu, ödeme planı olarak borçuların 0.598'nin ödeme planı olan olduğunu, ibra edilen olarak belirtilen borçuların 1.00'i ibra edilen olarak bulunduğu görülmüştür. Duyarlılık alanına bakıldığında borçuların doğru tespit edilme oranlarını inceleyecek olursak, tahsilat yapılamayan olarak belirtilen borçuların 1.00'ü, tahsilat yapılan olarak belirtilen borçuların 0.951'i, ödeme planı olarak borçuların 0.209'u, ibra edilen olarak belirtilen borçuların 1.00'i doğru tespit edildiğini görülmüştür. Karar ağacı sınıflandırma işlemi sonucunda verilerin sınıflara göre detaylı doğruluk bilgileri Çizelge 4.2'de verilmiştir.

Çizelge 4.2. C4.5 sınıflara göre detaylı doğruluk

	Doğru Pozitif Oranı	Yanlış Pozitif Oranı	Kesinlik	Duyarlılık	F ₁ Skoru	Class
	1.00	0.037	0.984	1.00	0.992	Tahsilat Yapılamayan
	0.951	0.065	0.807	0.951	0.873	Tahsilat Yapılan
	0.209	0.012	0.598	0.209	0.310	Ödeme Planı Olan
	1.00	0.00	1.00	1.00	1.00	İbra Edilen
Ağırlık Ortalaması	0.927	0.41	0.914	0.927	0.912	

C4.5 algoritmasının 34320 tane tahsilat yapılamayan borçlunun hepsini tahsilat yapılamayan olarak tahmin ettiği, 11023 tane tahsilat yapılan borçlunun 10478 tanesi tahsilat yapılan 545 tanesi ödeme planı olan olarak tahmin ettiği, 3876 tane ödeme planı olan borçlunun 566 tanesini tahsilat yapılamayan 2498 tanesini tahsilat yapılan ve 812 tanesini ödeme planı olan olarak tahmin ettiği, 292 tane ibra edilmiş olan borçlunun hepsini ise ibra edilmiş olarak tahmin ettiğini belirlenmiştir. Karar ağacı sınıflandırma işlemi sonucunda verilerin sınıflara göre karışıklık matrisi Çizelge 4.3'de verilmiştir.

Çizelge 4.3. C4.5 karışıklık matrisi

Tahsilat Yapılamayan	Tahsilat Yapılan	Ödeme Planı Olan	İbra Edilen	
34320	0	0	0	Tahsilat Yapılamayan
0	10478	545	0	Tahsilat Yapılan
566	2498	812	0	Ödeme Planı Olan
0	0	0	292	İbra Edilen

Yapılan inceleme sonucunda tahsilat yapılamayan ve ibra edilen borçluların hepsini doğru sınıflandırdığı gözlemlenmiştir. Tahsilat yapılan olan 11023 tane borçlunun ise 10478 tanesini doğru sınıflandırdığı görülürken iken ödeme planı olan borçlularda aynı verimlilikte sınıflandıramadığı gözlemlenmiştir.

C4.5 algoritmasıyla test için hazırlanan 100 satır veriyle karar ağacı tekrar test edilmiştir. 100 adet verinin 92 tanesini doğru sınıflandırır iken 8 tanesini yanlış sınıflandırdığı görülmüştür. Çizelge 4.4'de test verisine ait karışıklık matrisi verilmiştir.

Çizelge 4.4. C4.5 test verisi karışıklık matrisi

Tahsilat Yapılamayan	Tahsilat Yapılan	Ödeme Planı Olan	İbra Edilen	
74	0	0	0	Tahsilat Yapılamayan
0	15	7	0	Tahsilat Yapılan
1	0	3	0	Ödeme Planı Olan
0	0	0	0	İbra Edilen

Çizelge 4.4’de görüldüğü üzere, testin sonucunda 74 adet tahsilat yapılamayan borçlunun hepsini tahsilat yapılamayan, 22 adet tahsilat yapılan borçluyu 15 tanesini tahsilat yapılan geri kalanı ise ödeme planı olan borçlu, 4 adet ödeme planı olan borçlunun 3 tanesi ödeme planı olan 1 adet tahsilat yapılan olarak sınıflandırdığı gözlemlenmiştir.

Naive Bayes sınıflandırma işlemi sonucunda toplam 49511 tane veriden 44025 (%88.9196)’ının başarılı bir şekilde sınıflandırılmış olduğu 5486 (%11.0804)’inin ise doğru sınıflandırılmadığı görülmüştür. Ayrıca kök ortalama kare hatası 0.1987, görel mutlak hatası %25.8409, kök görel kare hatası %58.348 olduğu tespit edilmiştir. Naive Bayes sınıflandırma işlemi sonucunda verilerin sınıflandırılmasına yönelik özet bilgiler Çizelge 4.5’de verilmiştir.

4.2. Naive Bayes

Araştırmanın amaçlarına yönelik olarak verilerin analizini C4.5 karar ağacını kullanılarak elde edilen bulgular bu bölümde yer almaktadır.

Araştırmanın verilerini oluşturan borçluların, borçlarını geri ödeme eğilimlerini inceleyip karar ağacını oluşturabilmek amacıyla cinsiyet, yaş, doğum yeri, merkez bankası riski, alacaklı sayısı, anapara, toplam bakiye, toplam tahsilat, ikamet il, ikamet ilçe gibi nitelikleri Naive Bayes algoritmasında 10 kez çapraz doğrulama işlemi uygulanmış ve sınıflandırma işleminin detaylı sonuçları CD'deki EK-2'de verilmiştir. Toplam 49511 borçludan 44.025 (%88.9196) borçluyu başarılı bir şekilde sınıflandırılmış olduğunu, 5486 (%11.0804) borçluyu ise hatalı sınıflandırmış olduğu görülmektedir. Naive Bayes sınıflandırma işlemi sonucunda verilerin sınıflandırılmasına yönelik özet bilgiler Çizelge 4.5'de verilmiştir.

Çizelge 4.5. Naive Bayes sınıflandırma sonucu özeti

Sınıflandırma İstatistikleri	Değer	Oran (%)
Başarılı Sınıflandırma Örneklemesi	44025	88.9196
Hatalı Sınıflandırma Örneklemesi	5486	11.0804
Kappa İstatistiği	0.7607	
Ortalama mutlak hata	0.0599	
Kök ortalama kare hatası	0.1987	
Görelî mutlak hatası	25.8409 %	
Kök görelî kare hatası	58.348 %	
Toplam örnek sayısı	49511	

Naive Bayes algoritmasıyla 34320 tane tahsilat yapılamayan borçlunun 34298 tanesini tahsilat yapılamayan, 1 tanesini tahsilat yapılan ve 21 tanesini ibra olarak tahmin ettiği, 11023 tane tahsilat yapılan borçlunun 145 tanesini tahsilat yapılamayan, 7201 tanesi tahsilat yapılan, 3664 tanesi ödeme planı olan ve 13 tanesini ibra edilen olarak tahmin ettiği, 3876 tane ödeme planı olan borçlunun 595 tanesini tahsilat yapılamayan, 1035 tanesi tahsilat yapılan 2243 ödeme planı olan olarak tahmin ettiği,

292 tane ibra edilmiş olan borçlunun ise 6 tanesini tahsilat yapılan, 2 tanesini ödeme planı olan ve 283 tanesini ibra edilmiş olarak tahmin edildiği belirlenmiştir. Naive Bayes sınıflandırma işlemi sonucunda verilerin sınıflara göre karışıklık matrisi Çizelge 4.6'de verilmiştir.

Çizelge 4.6. Naive Bayes karışıklık matrisi

Tahsilat Yapılamayan	Tahsilat Yapılan	Ödeme Planı Olan	İbra Edilen	
34298	1	21	0	Tahsilat Yapılamayan
145	7201	3664	13	Tahsilat Yapılan
595	1035	2243	0	Ödeme Planı Olan
0	6	2	283	İbra Edilen

4.3. C4.5 ve Naive Bayes Bulgularının Karşılaştırılması

Bu bölümde, C4.5 ve Naive Bayes sınıflandırma teknikleriyle elde edilen sonuçlar karşılaştırılmıştır.

Çizelge 4.1 ve Çizelge 4.5 karşılaştırıldığında, başarılı sınıflandırılan örnek sayısı C4.5'de 45.856 iken Naive Bayes'de 44.025 olduğu yani örneklemede 1831 (%3.7911) kişiyi daha az başarılı bir şekilde sınıflandırdığı; hatalı sınıflandırılan örnek sayısı C4.5'de 3.609 iken Naive Bayes'de 5.486 olduğu yani örneklemede 1877 (%3.7911) kişiyi daha fazla hatalı bir şekilde sınıflandırdığı tespit edilmiştir. Kappa istatistiği tam anlaşma durumunda değeri 1 olur. Kappa istatistiği değeri 1'e ne kadar yakınsa anlaşma o derece iyidir. Hiç bir anlaşmanın olmaması durumunda kappa istatistiği değeri sifıra eşittir veya sıfırdan küçüktür. Buna göre 4.5 algoritmasında bu oran 0,8384 iken Naive Bayes ise 0,7607 oranına düştüğü tespit edilmiştir.

Çizelge 4.4 ve Çizelge 4.6 karşılaştırıldığında tahsilat yapılamayan, tahsilat yapılan, ödeme planı olan ve ibra edilen borçluların verilerinde farklılıklar olduğu belirlenmiştir. Tahsilat yapılamayan ve ibra edilen borçlular bakımından C4.5 algoritması hatasız bir şekilde sınıflandırma yapmasına karşın, Naive Bayes

algoritmasının tahsilat yapılamayan 34.320 borçludan 1 kişiyi tahsilat yapılan ve 21 kişiyi ise ödeme planı olan olarak sınıflandırdığı, ibra edilen 292 borçludan 6 kişiyi tahsilat yapılan ve 2 kişiyi ise ödeme planı olan olarak sınıflandırdığı görülmüştür. Tahsilat yapılan 11.023 borçlu bakımından C4.5 algoritması 10.478 kişiyi tahsilat yapılan ve 545 kişiyi ödeme planı olan olarak sınıflandırırken Naive Bayes algoritmasının tahsilat yapılan 7.201 borçludan 145 kişiyi tahsilat yapılamayan, 3.664 kişiyi ödeme planı olan ve 13 kişiyi ise ibra edilen olarak sınıflandırdığı tespit edilmiştir. Ödeme planı olan 3.876 borçludan C4.5 algoritmasında 566 kişinin tahsilat yapılamayan, 2.498 kişinin tahsilat yapılan ve 812 kişinin ödeme planı olan olarak sınıflandırıldığı, Naive Bayes algoritmasında ise 595 kişinin tahsilat yapılan, 1035 kişinin tahsilat yapılan ve 812 kişinin ödeme planı olan olarak sınıflandırdığı görülmüştür.

Dördüncü bölümde; araştırmanın amacına yönelik olarak kullanılan C4.5 ve Naive Bayes sınıflandırma teknikleri ile elde edilen bulgulara yer verilmiş, ardından her iki sınıflandırma tekniği ile elde edilen bulgular karşılaştırılmıştır.

5. TARTIŞMA

Varlık yönetim şirketlerine ait borçluların geri ödeme eğilimlerinin veri madenciliğinin sınıflandırma yöntemlerinden biri olan karar ağacı yönteminin kullanıldığı bu çalışmada borçlunun yaş, doğum yeri, cinsiyet gibi demografik bilgileri, anapara, toplam tahsilat, ikamet ettiği il, ikamet ettiği ilçe, ikamet ettiği mahalle gibi bilgileri dikkate alınarak veri madenciliğinde sınıflandırma algoritmaları yardımıyla incelenmiş; verilerin tekrar değerlendirilmesi sonrasında ülke gelir skoru, il gelir skoru, ilçe gelir skoru, kefil, pasif ve ödeme eğilimi algoritmalarının karar ağacı oluşturmada yeterli verimi sağlamadığı görüldüğünden başlangıçta 20 tane olan nitelikler bilgi kazanç nitelik değerlendirmesi (Information Gain Attributes Evaluation) algoritmasıyla değerlendirildikten sonra katkısı düşük olan nitelikler elendikten sonra 14'e düşürülmüştür. Bu işlemin uygulanmasında eşik değeri .0000455'den küçük olan nitelikler elenmiştir. Tosun (2006) tarafından yapılan çalışmada da karar ağaçlarındaki eşik değerinin yükseltilmesiyle kullanılan niteliklerin sayısının azalacağı sonucuna ulaşılmıştır.

Araştırmadaki 10 kategorik ve 10 nümerik değer olmak üzere toplamda 49.512 borçluya ait veriler incelenmiş ve verilerin analizi için karar ağacı oluşturmada C4.5 algoritması kullanılmıştır. WEKA programında da kullanılabilen J48 (C4.5 için geliştirilmiş ve aynı sonuçları veren algoritma) açık kaynak algoritmasının seçilmesinin sebebi verilerin niteliklerinin kategorik ve nümerik değerlerin bir arada değerlendirilmesine ihtiyaç duyulmasıdır. Bu bakımdan gerek veri sayısının fazla olması gerekse farklı nitelikleri bir arada bulundurması nedeniyle veri madenciliği araştırma verilerinin analizinde kullanılacak en etkili yöntem olduğu söylenebilir. Nitekim Çakır (2008) tarafından yapılan çalışmada veri sayısının ve veri kalitesinin, uygulamaların başarısında önemli birer faktör olduğu sonucuna ulaşılmıştır.

Araştırmada borçlularının cinsiyet, doğum yeri, yaş gibi demografik nitelikleri ile borç ödeme eğilimleri arasındaki ilişki incelenmiştir. Borçluların doğum yeri niteliği açısından cinsiyet, sigorta durumu, merkez bankası riski, ikamet edilen ilçe,

yaş gibi niteliklerle beraber borçlu kategorilerinin değiştiği ancak yalnızca doğum yeri Diyarbakır olan borçlular için ikamet edilen ilçe niteliğine bağlı olarak alt dal olduğu yani Diyarbakır'da doğan borçluların tamamının tahsilat yapılan sınıfında olduğu; borçluların yaş niteliği açısından yaş niteliğinin tek başına bir fark yaratmadığı ancak doğum yerleriyle birlikte incelediğinde doğum yerleri Adana, İzmir, Kadıköy, Şanlıurfa, Kahramanmaraş, Kütahya, Gaziantep, Kadirli, Eskişehir, Göle, Derinkuyu, Salihli, Kozan, Çarşamba, Afyonkarahisar, İzmit olan borçluların yaşa göre geri ödeme eğilimlerinin farklılık gösterdiği; borçluların cinsiyeti açısından Türkiye'nin birçok şehir ve ilçesinde cinsiyet ayrıştırıcı bir nitelik olmazken Gaziantep şehrinin Araban ilçesinde bu durumun farklı olduğu ve bu ilçede yaşayan kadınların, borçlarını geri ödediği görülürken erkeklerin ise ödeme planını kabul edip geri ödeme aşamasına geçemediği görülmüştür. Ceylan (2006) tarafından yapılan çalışmada tüketici kredisi müşterisinin gelir seviyesi, ev sahibi olup olmaması ve cinsiyetiyle tüketici kredisinin geri dönmemesi arasında anlamlı bir ilişki olduğu belirtilmiş iken incelenen diğer değişkenlerden tüketici kredisi müşterisinin yaşı, medeni durumu, öğrenim durumu, gelirinin kaynağı, kredi kartı sahibi olup olmaması ile kullanılan tüketici kredisinin mahiyeti, miktarı ve vadesi arasında anlamlı bir ilgiye rastlanılmadığı sonucuna ulaşılmıştır. Hörkkö (2010) tarafından yapılan çalışmada sosyodemografik ve davranışsal özelliklerin dikkate değer bir etki yarattığı ve bunların en önemli sosyodemografik nitelikleri gelir, yaş, eğitim ve milliyet olurken kredi kullanılabilirlik skoru, borcun büyüklüğü gibi davranışsal özelliklerinden önemli bir güce sahip olduğu sonucuna ulaşılmıştır. Yılmaz ve ark. (2013) tarafından yapılan çalışmada yaş, cinsiyet, medeni durum, pozisyon ve gelir düzeyi gibi demografik türde birçok değişkenle kredi kartını ilgilendiren kategorik değişkenler arasında çok sayıda anlamlı ilişkiler olduğunu sonucuna ulaşılmıştır. Amponsah (2016) tarafından yapılan çalışmada öğrencilerin borçlularının cinsiyetlerinin borç geri ödemelerini olumlu yönde etkilediği tespit edilmiş ve kız öğrenci borçlularının, erkek meslektaşlarına göre biraz daha fazla seviyede borcunu geri ödedikleri sonucuna ulaşılmıştır. Ramanujam and Vidya (2017) tarafından yapılan çalışmada cinsiyet, yaş, eğitim düzeyi, mesleki geçmiş deneyim yılı, gibi demografik özellikler ile kredi geri ödeme performansların arasında anlamlı bir fark oluşturmadığı sonucuna ulaşılmıştır.

Araştırmanın sonuçları, borçluların borçlarını geri ödeme eğilimlerinde; sigorta tek başına bir fark yaratmadığı ancak doğum yerleriyle birlikte incelediğinde

doğum İzmir ve Konya olan borçlularda sigorta bilgisine bağlı göre geri doğum yerleriyle birlikte incelediğinde doğum yeri Samsun, Ankara, Antalya, Akkuş ve Konya olan borçlularda alacaklı sayısı bilgisine bağlı göre geri ödeme eğilimlerinin farklılık gösterdiğini ortaya koymuştur. Ayrıca borçluların borçlarını geri ödeme eğilimlerinde; merkez bankası riski bilgisi tek başına bir fark yaratmadığı ancak doğum yerleriyle birlikte incelediğinde doğum yeri Ankara, Malatya, Manisa, Beykoz ve Kırıkkale olan borçlularda merkez bankası riski bilgisine bağlı göre ödeme eğilimlerinin farklılık gösterdiği tespit edilmiştir. Bir diğer sonuç ise borçluların, borçlarını geri ödeme eğilimlerinde anapara bilgisinin etkisi incelendiğinde karar ağacında toplam bakiye ve doğum yeri bilgisiyle beraber borçlunun geri ödeme eğilimlerinin farklılık gösterdiği, ibra edilen borçluların karar ağacının tek bir noktasında görüldüğü, borçlunun demografik ve lokasyon bilgilerinin bu durumu etkilemediğini göstermektedir. Bu durum borçluların niteliklerinin, borçları geri ödeme eğilimlerinde farklılıklar oluşturduğu şeklinde ifade edilebilir. Bu açıdan incelendiğinde bankaların kredi verirken, borçluların sigortası, merkez bankası riski ve anapara bilgisi gibi niteliklerini dikkate alması verilen kredilerin geri ödenememe riskini düşürebileceği düşünülebilir. Bilen (2009) tarafından yapılan çalışmada doğru insanın doğru özellikler ile doğru yerde kullanılması ve doğru hedeflere odaklanması ile ciddi anlamda bir fayda sağlanacak, gerek çalışanlar gerekse de işletme için verimlilik ve etkinliğin artacağı sonucuna ulaşılmıştır. Bu araştırmanın sonuçları, benzer şekilde müşteri seçiminde ve sınıflandırılmasında bankalar gibi kredi veren sektörler tarafından geliştirilecek ve veri madenciliğinin etkili bir şekilde kullanılacağı bir sistemin oluşturulmasında kullanılabileceği şeklinde yorumlanabilir.

Beşinci bölümde; araştırmada elde edilen bulgular ile literatürde yer alan ilgili araştırmaların bulguları karşılaştırılmıştır.

6. SONUÇLAR

Günümüzde teknolojinin de gelişmesiyle birlikte büyük miktardaki verinin daha kolay elde edilebilmesinden dolayı “bilgi çıkarımı” oldukça önem kazanmıştır. Bu durum verilerin yararlı bilgilere dönüştürülebilme ihtiyacını beraberinde getirmiş böylece veri madenciliği gün geçtikçe dikkatleri üzerinde toplamaya başlamıştır. Başta bankacılık, finans ve pazarlama olmak üzere sağlık, insan kaynakları, telekomünikasyon, sigortacılık gibi pek çok alanda veri madenciliği veri elde edilmesinde kullanılan en önemli kaynaklardan biri haline gelmiştir. Bu bakımdan veri madenciliği, çok sayıda borçluluyla çalışmakta olan varlık yönetimi şirketlerinin borçlularının geri ödeme eğilimlerinin analizinde büyük önem arz etmektedir.

Bu çalışmada, varlık yönetimi şirketlerine ait borçluların geri ödeme durumları değerlendirilmiş, sınıflandırma yöntemlerinden karar ağacı ve Naive Bayes ile borçluların geri ödeme eğilimleri incelenmiş ve mevcut niteliklerin sürece katmış olduğu farklı anlamlar keşfedilmeye çalışılmıştır. Borçlunun yaş, doğum yeri, cinsiyet gibi demografik bilgileri, anapara, toplam tahsilat, ikamet ettiği il, ikamet ettiği ilçe, ikamet ettiği mahalle gibi bilgileri dikkate alınarak veri madenciliğinde sınıflandırma algoritmaları yapılmış, verilerin her birinin kendi arasındaki bağlantısı incelenmiş ve elde edilen bulgular yorumlanmıştır. Başlangıçta 21 tane olan nitelikler bilgi kazanç nitelik değerlendirmesi (Informaiton Gain Attributes Evaluation) algoritmasıyla değerlendirildikten sonra katkısı düşük olan nitelikler elendikten sonra 15’ e düşürülmüştür. Verilerin tekrar değerlendirilmesi sonrasında ülke gelir skoru, il gelir skoru, ilçe gelir skoru, kefil, pasif ve ödeme eğilimi algoritmalarının karar ağacı oluşturmada yeterli verimi sağlamadığı görülmüştür. Verilerin analizinde C4.5 karar ağacı ve Naive Bayes algoritması ile veriler değerlendirilmiş, elde edilen sonuçlar karşılaştırılmıştır.

Araştırmanın sonuçları C4.5 karar ağacı bakımından incelendiğinde; karar ağacında kök düğümün toplam tahsilat olduğu ve dalların buradan başladığını göstermektedir. Kök düğümün altında toplam bakiye, anapara, doğum yeri, ikamet

ettiği il, ikamet ettiği ilçe ve ikamet ettiği mahalleye göre farklı dallar oluşmuştur. Araştırmanın bulgularından elde edilen sonuçlar şu şekildedir:

- Borçluların, borçlarını geri ödeme eğilimlerinde doğum yeri bilgilerinin önemi incelendiğinde; cinsiyet, sigorta durumu, merkez bankası riski, ikamet edilen ilçe, yaş gibi niteliklerle beraber borçlu kategorilerinin değiştiği ancak yalnızca doğum yeri Diyarbakır olan borçlular için ikamet edilen ilçe niteliğine bağlı olarak alt dal oluştuğu yani Diyarbakır'da doğan borçluların tamamının tahsilat yapılan sınıfında olduğu görülmüştür. Bu durum kişilerin doğduğu yere bağlı olarak çevresiyle ve akrabalarıyla olan etkileşim düzeylerinin farklılık gösterebileceği, ülkemizin doğusunda yer alan Diyarbakır gibi illerde ailelerin daha kalabalık olmasından dolayı borçlulara ailesi veya akrabaları tarafından destek verilmiş olabileceği şeklinde düşünülebilir.
- Borçluların, borçlarını geri ödeme eğilimlerinde yaş bilgilerinin önemi incelendiğinde tek başına bir fark yaratmadığı ancak doğum yerleriyle birlikte incelendiğinde doğum yerleri Adana, İzmir, Kadıköy, Şanlıurfa, Kahramanmaraş, Kütahya, Gaziantep, Kadirli, Eskişehir, Göle, Derinkuyu, Salihli, Kozan, Çarşamba, Afyonkarahisar, İzmit olan borçluların yaşa göre geri ödeme eğilimlerinin farklılık gösterdiği görülmüştür. Bu durum elde edilen bulgular doğrultusunda doğum yeri ve yaş bilgisinin birlikte değerlendirilmesi gerektiği şeklinde yorumlanabilir.
- Borçluların, borçlarını geri ödeme eğilimlerinde cinsiyet bilgilerinin önemi incelendiğinde; Türkiye'nin birçok şehir ve ilçesinde cinsiyet ayrıştırıcı bir nitelik olmazken Gaziantep şehrinin Araban ilçesinde bu durumun farklı olduğu ve bu ilçede yaşayan kadınların, borçlarını geri ödediği görülürken erkeklerin ise ödeme planını kabul edip geri ödeme aşamasına geçemediği sonucuna ulaşılmıştır.
- Borçluların, borçlarını geri ödeme eğilimlerinde sigortanın tek başına bir fark yaratmadığı ancak doğum yerleriyle birlikte incelendiğinde; doğum yeri İzmir ve Konya olan borçlularda sigorta bilgisine bağlı göre geri ödeme eğilimlerinin farklılık gösterdiği görülmüştür.
- Borçluların, borçlarını geri ödeme eğilimlerinde sigorta bilgisinin tek başına bir fark yaratmadığı ancak doğum yerleriyle birlikte; doğum yeri Samsun,

Ankara, Antalya, Akkus ve Konya olan borçlularda alacaklı sayısı bilgisine bağlı göre geri ödeme eğilimlerinin farklılık gösterdiği görülmüştür.

- Borçluların, borçlarını geri ödeme eğilimlerinde merkez bankası riski bilgisinin tek başına bir fark yaratmadığı ancak doğum yerleriyle birlikte incelendiğinde; doğum yeri Ankara, Malatya, Manisa, Beykoz ve Kırıkkale olan borçlularda merkez bankası riski bilgisine bağlı göre geri ödeme eğilimlerinin farklılık gösterdiği görülmüştür.
- Borçluların, borçlarını geri ödeme eğilimlerinde anapara bilgisinin etkisi incelendiğinde karar ağacında toplam bakiye ve doğum yeri bilgisiyle beraber borçlunun geri ödeme eğilimlerinin farklılık gösterdiği görülmüştür.
- İbra edilen borçluların karar ağacının tek bir noktasında görüldüğü, borçlunun demografik nitelikleri ve lokasyon nitelikleri bu durumu etkilemediği görülmüştür.

Araştırmada kullanılan karar ağacı ve Naive Bayes sınıflandırma yöntemleri ile elde edilen bulgular karşılaştırıldığında elde edilen sonuçlar şu şekildedir:

- Başarılı sınıflandırılan örnek sayısı C4.5'de 45.856 iken Naive Bayes'de 44.025 olduğu yani örneklemede 1831 (%3.7911) kişiyi daha az başarılı bir şekilde sınıflandırdığı; hatalı sınıflandırılan örnek sayısı C4.5'de 3.609 iken Naive Bayes'de 5.486 olduğu yani örneklemede 1877 (%3.7911) kişiyi daha fazla hatalı bir şekilde sınıflandırdığı görülmüştür.
- Naive Bayes algoritmasıyla 34320 tane tahsilat yapılamayan borçlunun 34298 tanesini tahsilat yapılamayan, 1 tanesini tahsilat yapılan ve 21 tanesini ibra olarak tahmin ettiği, 11023 tane tahsilat yapılan borçlunun 145 tanesini tahsilat yapılamayan, 7201 tanesini tahsilat yapılan, 3664 tanesini ödeme planı olan ve 13 tanesini ibra edilen olarak tahmin ettiği, 3876 tane ödeme planı olan borçlunun 595 tanesini tahsilat yapılamayan, 1035 tanesini tahsilat yapılan 2243 ödeme planı olan olarak tahmin ettiği, 292 tane ibra edilmiş olan borçlunun ise 6 tanesini tahsilat yapılan, 2 tanesini ödeme planı olan ve 283 tanesini ibra edilmiş olarak tahmin edildiği görülmüştür.
- Tahsilat yapılan, ödeme planı olan ve ibra edilen borçluların verilerinde farklılıklar olduğu belirlenmiştir. Tahsilat yapılamayan ve ibra edilen borçlular bakımından C4.5 algoritması hatasız bir şekilde sınıflandırma yapmasına karşın, Naive Bayes algoritmasının tahsilat yapılamayan 34.320

borçludan 1 kişiyi tahsilat yapılan ve 21 kişiyi ise ödeme planı olan olarak sınıflandırdığı, ibra edilen 292 borçludan 6 kişiyi tahsilat yapılan ve 2 kişiyi ise ödeme planı olan olarak sınıflandırdığı görülmüştür. Tahsilat yapılan 11.023 borçlu bakımından C4.5 algoritması 10.478 kişiyi tahsilat yapılan ve 545 kişiyi ödeme planı olan olarak sınıflandırırken Naive Bayes algoritmasının tahsilat yapılan 7.201 borçludan 145 kişiyi tahsilat yapılamayan, 3.664 kişiyi ödeme planı olan ve 13 kişiyi ise ibra edilen olarak sınıflandırdığı tespit edilmiştir. Ödeme planı olan 3.876 borçludan C4.5 algoritmasında 566 kişinin tahsilat yapılamayan, 2.498 kişinin tahsilat yapılan ve 812 kişinin ödeme planı olan olarak sınıflandırıldığı, Naive Bayes algoritmasında ise 595 kişinin tahsilat yapılan, 1035 kişinin tahsilat yapılan ve 812 kişinin ödeme planı olan olarak sınıflandırdığı görülmüştür.

Veri madenciliği uygulaması neticesinde borçluların geri ödeme eğilimleri C4.5 ve Naive Bayes algoritmaları kullanılarak değerlendirilmiştir. Araştırmanın sonuçları borçluların kategorilerini belirlemede C4.5 algoritmasının Naive Bayes algoritmasına oranla daha başarılı olduğunu göstermektedir. Ayrıca Naive Bayes algoritması ile yalnızca borçlularının olasılıksal olarak sınıflandırılması yapılabilmesine karşın C4.5 algoritması sonucunda elde edilen karar ağacıyla hem borçluların sınıflandırılması yapılmış hem de borçlunun borcu geri ödeme eğilimine yönelik kurallar oluşturulmuştur. C4.5 algoritması ile elde edilen bu kurallar dikkate alınarak, bir borçlunun özelliklerine göre geri ödeme eğiliminin ne olduğu öngörülebilecektir. Bu sayede, borçlulara sunulan geri ödeme planlarının daha nitelikli hale getirilebileceği ve varlık yönetimi şirketlerinin belirledikleri hedefe daha etkili bir şekilde ulaşmalarına yardımcı olacağı düşünülebilir. Sonuç olarak, aranılacak özelliklerin ve izlenilecek yolların düzgün bir şekilde belirlenmesinden sonra varlık yönetimi şirketi açısından, kendi çalışanlarından beklediği performansta bir fayda sağlanmış, gerek çalışanlar gerekse de işletme için verimlilik ve etkinlik artmış olacaktır. Tez kapsamında gerçekleştirilen çalışma sonuçları, veri madenciliğinin varlık yönetimi şirketlerindeki borçlunun geri ödeme eğilimlerinin analizinde uygulanabilirliğini gösterdiği şeklinde söylenebilir.

Altıncı bölümde; araştırmanın amacına yönelik olarak elde edilen bulgulara yer verilmiştir.

7. ÖNERİLER

Elde edilen bulgularda izlenen yöntemler incelenerek, farklı niteliklerle süreç yeniden değerlendirilebilir. Özellikle borçlunun malvarlıkları var veya yok olma durumu yerine gayrimenkul, menkul, araç ve maaş gibi farklı kategorilerde incelenerek, bunların bir kısmına sahip olan borçluların borcunu geri ödeme eğilimi analiz edilebilir. Yapılan analiz sonucunda elde edilen sınıflandırma kurallarıyla varlık yönetimi firmasına tavsiye niteliğinden sunulmuş ve borcun geri ödeme eğiliminde hangi özelliklerin daha etkili olduğu belirterek verimliliğin artırılmasında katkı sağlanabilir.

Araştırmada sınıflandırma tekniklerinden C4.5 ve Naive Bayes teknikleri kullanılmıştır. Borçların geri ödeme eğilimi analizinde diğer sınıflandırma teknikleri kullanılarak farklı araştırmalar yapılabilir.

Araştırmada varlık yönetim şirketlerinin borçlularının geri ödeme eğilimleri incelenmiştir. Borçlularla doğrudan ilgili olan finans sektörlerindeki şirketlerin borçlularının geri ödeme eğilimlerine yönelik çalışmalar yapılabilir.

8. KAYNAKLAR

Agrawa, G. L., Gupta, H. (2013). Optimization of C4.5 Decision Tree Algorithm for Data Mining Application, *International Journal of Emerging Technology and Advanced Engineering*, 3(3):341-345.

Ak, A. (2014). Meşruiyet Analizi Bağlamında Varlık Yönetim Şirketleri, *İslam Hukuku Araştırmaları Dergisi*, 24:127-148.

Akçay, A. (2014). Bilgi ve Belge Yönetiminde Veri Madenciliği. İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, İstanbul, (Danışman: Doç.Dr. Ümit Konya).

Akpınar, H. (2000). Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği. İ.Ü İşletme Fakültesi Dergisi, 29(1):1-22.

Aksu, M. Ç. Ve Karaman, E. (2017). Karar Ağaçları İle Bir Web Sitesinde Link Analizi ve Tespiti, *Acta Infologica*, 1(2): 84-91.

Alan, M. A. (2014). Karar Ağaçlarıyla Öğrenci Verilerinin Sınıflandırılması. Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 28(4):101-112.

Amponsah, A. M. (2016). Students' Indebtedness and Their Repayments: The Case of Ghana Technology University College.

Arpacı, S. A. ve Kalıpsız, O. (2018). Yazılım Hata Sınıflandırmasında Farklı Naive Bayes Tekniklerinin Kıyaslanması, *Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 7(1):1-13.

Atalar, M. ve Şahin, N. K. (2013). Kredi Tahsisinde Rating Yönteminin Sorunlu Krediyeye Etkisi, *Balıkesir Üniversitesi, Sosyal Bilimler Enstitüsü Dergisi*, 16(30):89-113.

Aydın, S. ve Özkul A. E. (2015). Veri Madenciliği ve Anadolu Üniversitesi Açıköğretim Sisteminde Bir Uygulama. *Eğitim ve Öğretim Araştırmaları Dergisi*, 4(3):36-44.

Bahadır, C. (2018). Sorunlu Kredilerin Yönetiminde Varlık Yönetim Şirketlerinin Rolü ve Türkiye Uygulaması. *Başkent Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, Ankara, (Danışman: Doç.Dr. Adalet Hazar).*

Bırtıl, F. S. (2011). Kız Meslek Lisesi Öğrencilerinin Akademik Başarısızlık Nedenlerinin Veri Madenciliği Tekniği İle Analizi. *Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Afyonkarahisar, (Danışman: Yrd.Doç.Dr. Ömer Deperlioğlu).*

Bilen, H. (2009). Bankacılık Sektöründe Personel Seçimi ve Performans Değerlendirilmesine İlişkin Veri Madenciliği Uygulaması Analizi, *Gazi Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara, (Danışman: Prof.Dr. Ertan Güner).*

- Budak, H. (2018). Özellik Seçim Yöntemleri ve Yeni Bir Yaklaşım. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 22(Özel Sayı):21-31.
- Ceylan, O. (2006). Türkiye’de Yıllara Göre Kredi Kartı Kullanım Durumu ve Batık Kredi Kart Borç Sahiplerinin Profili. İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, İstanbul, (Danışman: Prof.Dr. Ozan Ceylan).
- Çakır, Ö. (2008). Veri Madenciliğinde Sınıflandırma Yöntemlerinin karşılaştırılması “Bankacılık Müşteri Veritabanı Üzerinde bir Uygulama”. Marmara Üniversitesi Sosyal Bilimler Enstitüsü, Doktora Tezi, İstanbul, (Danışman: Prof.Dr. İsmail Hakkı Armutlulu).
- Çalış, A., Kayapınar S. ve Çetinyokuş T. (2014). Veri Madenciliğinde Karar Ağacı Algoritmaları ile Bilgisayar ve İnternet Güvenliği Üzerine Bir Uygulama, Endüstri Mühendisliği Dergisi, 25(3-4):2-19.
- Dunham, M. H. (2003). Data Mining Introductory and Advanced Topics. New Jersey: Pearson Education, Inc.
- Deveci, M. A. (2018). Müşteri İlişkileri Yönetiminde Veri Madenciliği ve İş Zekası Uygulamaları. Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, Sivas, (Danışman: Doç.Dr. Oğuz Kaynar).
- Erduran, G. Y. (2017). Online Müşteri Şikayetlerinin Veri Madenciliği İle İncelenmesi. Trakya Üniversitesi Sosyal Bilimler Enstitüsü, Doktora Tezi, Edirne, (Danışman: Doç.Dr. Fatma Lorcü).
- Ersöz, T., Merdin, D. ve Ersöz F. (2015). Veri Madenciliği Yöntemi İle Memnuniyet Algısının Araştırılması, ISITES2015 Valencia, Spain.
- Flach, P. ve Kull, M. (2015). Precision-Recall-Gain Curves: PR Analysis Done Right. Advances in Neural Information Processing Systems 28, 4(1).
- Gökgöz, K. B. (2015). Veri Madenciliği İle Tıbbi Cihaz Bakım Karar Modeli. Başkent Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, Ankara, (Danışman: Prof.Dr. Nermin Özgülbaş; Doç.Dr. Ali Serhan Koyuncugil).
- Güldal, H. ve Çakıcı, Y. (2017). Eğitsel Veri Madenciliği. Balkan Eğitim Araştırmaları 2017 / Balkan Educational Studies 2017 1:135-143.
- Han, J., Kamber, M.,” Data Mining: Concept and Techniques”, Morgan Kaufmann Publishers, 2000.
- Hand, D., Mannila, H., Smyth, P.,“ Principles of Data Mining“, Cambridge, CA: MIT Press, 2001.
- Hörkkö, M. (2010). The Determinants of Default in Consumer Credit Market (Master Degree Thesis). Department of Accounting and Finance Aalto University.
- İnan, O., (2003). Veri Madenciliği. Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Konya, (Danışman: Doç.Dr. Nihat Yılmaz).
- Kalikov, A., (2006), Veri Madenciliği ve Bir E-Ticaret Uygulaması. Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara, (Danışman: Yrd.Doç.Dr. Nursal Arıcı).
- Koç, S. (2010). Finansal Yeniden Yapılandırma: İstanbul ve Anadolu Yaklaşımları, Sonuçları, C. Ü. İktisadi ve İdari Bilimler Dergisi, 11:35-55.

- Maimon, O., Lior, Rokach. Data Mining and Knowledge Discovery Handbook, Springer, 2.bs., 2010.
- Özkan, Y. (2008), Veri Madenciliği Yöntemleri. Papatya Yayınları, İstanbul.
- Ramanujam, V. ve Arun Vidya, K. (2017). A Study on The Credit Repayment Behavior of Borrowers. International Research Journal of Business and Management – IRJBM, 10(8):9-18.
- Savaş, S., Topaloğlu, N. ve Yılmaz, M. (2012). Veri Madenciliği ve Türkiye’deki Uygulama Örnekleri. İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 11(21):1-23.
- Sharma, S., Agrawal, J. Ve Sharma, S. (2013). Classification Through Machine Learning Technique: C4.5 Algorithm Based on Various Entropies, International Journal of Computer Applications, 82(16):20-27.
- Seema, Rathi, M. ve Mamta (2012). Decision Tree: Data Mining Techniques. International Journal of Latest Trends in Engineering and Technology, 1(3):150-155.
- Tahtakılıç, A. K. (2004). Bankacılık Krizlerinde Varlık Yönetim Şirketleri ve Türkiye Üzerine Değerlendirmeler (Yeterlik Etüdü), Sermaye Piyasası Kurulu Ortaklıklar Finansman Dairesi.
- Tekten, E. (2018). Türk Ticaret Kanunundaki Hükümler Karşısında Varlık Yönetim Şirketlerinin Kuruluşu ve Kuruluşunun İzne Tabi Olması. Erciyes Üniversitesi Hukuk Fakültesi Dergisi, 13(1):183-216.
- Ture, E. (2015). Bankalarda Kredilerin Yönetimi ve Sorunlu Kredilerin Yeniden Yapılandırma Süreci. Marmara Üniversitesi Bankacılık ve Sigortacılık Enstitüsü, Yüksek Lisans Tezi, İstanbul, (Danışman: Yrd.Doç.Dr. Hayati Eriş).
- Tosun, T. (2006). Veri Madenciliği Teknikleriyle Kredi Kartlarında Müşteri Kaybetme Analizi. İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstanbul, (Danışman: Dr. Sabih Atadan).
- Varlık Yönetim Şirketlerinin Kuruluş Ve Faaliyet Esasları Hakkında Yönetmelik (2006). T.C. Resmi Gazete (26333, 1 Kasım 2006).
- Vikramkumar. Vijaykumar, B., Trilochan (2014). Bayes and Naive Bayes Classifier. arXiv.
- Yaşar, H. (2017). Standart Dışı Verilerin Paralel Programlama İle Analizi ve Modellenmesi. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Isparta, (Danışman: Yrd.Doç.Dr. Mehmet Albayrak).
- Yıldırım, E. (2007). Bankalarda Ticari Kredilendirme Süreci; Karşılaşılan Sorunlar ve Çözüm Yolları ile İlgili Ampirik Bir Çalışma. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü.
- Yılmaz, H., Sancar Budak, G. ve Başaran, B. (2013). Kredi Kartı Kullanım Alışkanlıklarında Kategorik Değişkenler ve Bireylerin Davranışsal Eğilimleri: Bilecik Örneği. Uluslararası Yönetim İktisat ve İşletme Dergisi, 9(19):31-49.
- Zhang, H. (2004). The Optimality of Naive Bayes, Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, Miami Beach, Florida, USA.

9. ÖZGEÇMİŞ

1990 yılında Samsun'da doğdu. İlkokulu Ticaret ve Sanayi Odası İlk Öğretim Okulunda okuduktan sonra Samsun Mithat Paşa Lisesi Fen-Matematik bölümünden 2017 yılında mezun oldu. 2008 yılında başladığı Sakarya Üniversitesi Bilgisayar Mühendisliği bölümünden 2013 yılında mezun oldu. 2011 yılından itibaren Yazılım Geliştiricisi ve Yazılım Geliştirme Uzmanı görevlerini üstlenerek çeşitli firmalarda çalıştı. Aynı zamanda 2014 yılında kayıt olduğu Haliç Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği bölümünde yüksek lisans öğrencisi olarak öğrenim hayatına devam etmektedir.