



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

**BANKACILIKTA DOLANDIRICILIK
YÖNTEMLERİ ve TESPİTİ**

YÜKSEK LİSANS TEZİ

**Hazırlayan
Çağrı CİVELEK**

**Danışman
Prof. Dr. Mübariz EMİNLİ**

İstanbul – 2019

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

**BANKACILIKTA DOLANDIRICILIK
YÖNTEMLERİ ve TESPİTİ**

YÜKSEK LİSANS TEZİ

**Hazırlayan
Çağrı CİVELEK**

**Danışman
Prof. Dr. Mübariz EMİNLİ**

İstanbul – 2019

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Programı Öğrencisi Çağrı CİVELEK tarafından hazırlanan “*Bankacılıkta Dolandırıcılık Yöntemleri ve Tespiti*” konulu çalışması jürimizce Yüksek Lisans olarak kabul edilmiştir.

Tez Savunma Tarihi: 27.06.2019

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

İmzası

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

Jüri Üyesi : Dr.Öğr.Üyesi Ülviye HACIZADE
:Haliç Üniversitesi

Jüri Üyesi : Dr.Öğr.Üyesi Faruk BULUT
: İstanbul Rumeli 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

Çağrı Civelek

ORIJINALLIK RAPORU

%**4**

BENZERLIK ENDEKSİ

%**1**

İNTERNET
KAYNAKLARI

%**0**

YAYINLAR

%**3**

ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1

Submitted to Haliç Üniversitesi

Öğrenci Ödevi

%**1**

2

www.tbb.org.tr

İnternet Kaynağı

%**1**

3

Submitted to TechKnowledge Turkey

Öğrenci Ödevi

%**1**

4

Submitted to Beykent Üniversitesi

Öğrenci Ödevi

<%**1**

5

Submitted to Istanbul Bilgi University

Öğrenci Ödevi

<%**1**

6

Submitted to Kadir Has University

Öğrenci Ödevi

<%**1**

7

Submitted to Afyon Kocatepe University

Öğrenci Ödevi

<%**1**

8

Submitted to Fırat Üniversitesi

Öğrenci Ödevi

<%**1**

9

Submitted to Marmara University

Öğrenci Ödevi

<%**1**

Görülür NWA

-
- 10 ATEŞ, Yalçın and KARABATAK, Murat. "Nicel Birliktelik Kuralları İçin Çoklu Minimum Destek Değeri", Fırat Üniv. Fen Bilimleri Enst., 2017.
Yayın <% 1
-
- 11 Submitted to University of Bath
Öğrenci Ödevi <% 1
-
- 12 www.researchprotocols.org
İnternet Kaynağı <% 1
-
- 13 Submitted to Higher Education Commission
Pakistan
Öğrenci Ödevi <% 1
-
- 14 www.dtic.mil
İnternet Kaynağı <% 1
-
- 15 destek.kmk.net.tr
İnternet Kaynağı <% 1
-
- 16 Q.A. Al-Gubri, M.A.M. Ariff, I.S. Saeh.
"Performance analysis of machine learning algorithms for power system dynamic security assessment", 4th IET Clean Energy and Technology Conference (CEAT 2016), 2016
Yayın <% 1
-
- 17 www.zmo.org.tr
İnternet Kaynağı <% 1
-
- 18 Submitted to Ondokuz Mayıs Üniversitesi
Öğrenci Ödevi <% 1
-

nr

19

acikerisim.ticaret.edu.tr

İnternet Kaynağı

<%1

20

www.slideshare.net

İnternet Kaynağı

<%1

21

polen.itu.edu.tr

İnternet Kaynağı

<%1

Alıntılarını çıkart

Kapat

Eşleşmeleri çıkar

< 5 words

Bibliyografyayı Çıkart

üzerinde

23/07/2019

TEZ ETİK BEYANI

Yüksek Lisans Tezi olarak sunduğum “Bankacılıkta Dolandırıcılık Yöntemleri ve Tespiti” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof.Dr Mübariz EMİNLİ ‘nin sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, 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.

ÇAĞRI CİVELEK

ÖNSÖZ

Günümüzde küreselleşme kavramının gelişimi, insanları daha hızlı ve rahat işlem yapabilmeye sevk etmiştir. Bu da küresel ölçekte bütün işlemlerin internet ve sanal ortamla buluşması demektir. Daha hızlı ve güvenli işlem, kolay kullanım, konum bağımsız olarak hizmet edebilme yeteneği ve kullanıcı portföyünün gelişimi gibi olumlu etkiler düşünüldüğünde bankacılık ve finans sektörü de sisteme ayak uydurmaya ve internet dünyasına girmeye karar vermiştir. Aslında bu tam anlamıyla bir karar olmaktan ziyade bir zorunluluk olmuştur.

Bankacılık ve finans kavramının sanal dünya ile buluşması bir takım tehditleri ve bu tehditlere bağlı olarak önlemleri de beraberinde getirmektedir. Doğrudan paranın bulunduğu bir ortam dolandırıcıların ve hırsızların dikkatini çekmektedir. Bütün bu yapı ve tehditler düşünülerek yazımız temel olarak bizleri korkutan dolandırıcılıkların tespiti, analizi ve kurumlar açısından değerlendirim gibi konuları örnek senaryolarla ifade etmiştir.

Çalışmamızda bizlerden desteklerini esirgemeyen, bizleri geliştirmeye, araştırmaya ve her zaman öğrenmeye iten saygıdeğer hocamız Sayın Prof. Dr. MÜBARİZ EMİNLİ ve HALIÇ ÜNİVERSİTESİ ailesine sonsuz teşekkürlerimi iletiyorum.

Bütün hayatımız boyunca her daim yolumuzu aydınlatan, desteklerini ve öğütlerini bizlerden esirgemeyen bizleri Ulu Önder MUSTAFA KEMAL ATATÜRK ilke ve inkılapları ile yetiştiren, eğitimin ışığını aşılayan değerli aileme sevgi, saygı ve şükranlarımı sunarım.

İstanbul, 2019

ÇAĞRI CİVELEK

İÇİNDEKİLER

Sayfa No.

TEZ ETİK BEYANI	iii
ÖNSÖZ	iv
İÇİNDEKİLER	I
KISALTMALAR	IV
ŞEKİLLER	V
ÖZET	VII
ABSTRACT	VIII
1. GİRİŞ	1
2. DÜNYADA ve TÜRKİYE’DEKİ BANKACILIK GELİŞİMİ	5
3. İNTERNET BANKACILIĞI	6
3.1. İnternet Bankacılığı İşlemleri.....	6
3.2. İnternet Bankacılığı Avantaj ve Dezavantajları	7
4. SOSYAL MÜHENDİSLİK	9
4.1. Yaşam Döngüsü	10
4.2. Kullanılan Genel Yöntemler	10
4.2.1. Temeli Teknoloji Olan Sosyal Mühendislik	11
4.2.2. Temeli İnsan Olan Sosyal Mühendislik	11
4.3. Şirket Üzeri Etkiler	12
5. DOLANDIRICILIK VE TÜRLERİ	13
5.1. İnternet Dolandırıcılığı	13
5.1.1. Bilgilerin İnternet Üzerinden Düzenlenmesi	14
5.1.2. İnternet Güvenliği İçin Neler Yapılmalı	15
5.2. Çağrı Merkezli Dolandırıcılıklar	16
5.2.1. Dolandırıcı Tarafından Çağrı Merkezinin Aranması	17

5.3. Kartlı Sistem Dolandırıcılığı	17
5.3.1. ATM Dolandırıcılığı	17
5.3.2. Kart Dolandırıcılığı	18
5.3.3. Kart Dolandırıcılığından Korunma	18
6. BANKACILIKTA DOLANDIRICILIK TESPİT YÖNTEMLERİ.....	19
7. VERİ MADENCİLİĞİ	21
7.1. Veri Madenciliği Tanımı	21
7.2. Veri Madenciliğinin Tarihçesi	23
7.3. Veri Madenciliği Süreçleri	24
7.4. Veri Madenciliği Modelleri.....	26
7.4.1. Tanımlayıcı Modeller.....	26
7.4.1.1. Kümeleme	26
7.4.1.2. Birliktelik Kuralları.....	27
7.4.2. Tahmin edici Modeller.....	28
7.4.2.1. Regresyon.....	28
8. SINIFLANDIRMA YÖNTEMİ	29
8.1. En Yakın k - Komşu Algoritması ile Sınıflandırma.....	29
8.2. Yapay Sinir Ağları ile Sınıflandırma	30
8.3. Karar Ağaçları ile Sınıflandırma	31
9. GEREÇ ve YÖNTEM.....	33
9.1. C4.5 Algoritması	33
9.2. Weka Yazılımı	35
9.2.1. Weka 'da Kullanılan Ekranlar.....	35
9.3. PHP	41
9.4. HTML	42
9.5. CSS.....	42
9.6. MySQL.....	43
10. VERİ SETİNİN HAZIRLANMASI	45
11. BULGULAR.....	51

12. UYGULAMA ARAYÜZLERİ.....	59
12.1. Uygulama Giriş Sayfası Arayüzü.....	59
12.2. Uygulama Anasayfa Arayüzü	60
12.3. Uygulama Sanal Pos Arayüzü.....	60
12.4. Uygulama Müşteri Risk Grubu Arayüzü	61
12.5. Uygulama Müşteri İşlemleri Arayüzü.....	61
12.6. Uygulama Müşteri Ekle Arayüzü.....	62
12.7. Uygulama Müşteri Güncelle Arayüzü	63
12.8. Uygulama Müşteri Sil Arayüzü	64
12.9. Uygulama Alışveriş Listeleme Arayüzü	64
13. TARTIŞMA	66
14. SONUÇ.....	68
15. KAYNAKLAR	70
16. ÖZGEÇMİŞ.....	72

KISALTMALAR

ARFF	: Attribute Relationship File Format
ATM	: Automatic Teller Machine
CSS	: Cascading Style Sheets
CSV	: Comma Separated Values
EFT	: Elektronik Fon Transferi
FN	: False Negative
FP	: False Positive
HTML	: Hyper Text Markup Language
MitM	: Man in The Middle Attack
PC	: Personel Computer
PHP	: Hypertext Preprocessor
POS	: Point of Sales Terminal
SMS	: Short Message Service
TN	: True Negative
TP	: True Positive
YSA	: Yapay Sinir Ağları
WEKA	: Waikato Environment for Knowledge Analysis

ŞEKİLLER

	Sayfa No.
Şekil 4.1. Sosyal mühendislik yaşam döngüsü	10
Şekil 5.1. Dolandırıcılık yapan çetenin organizasyonu	14
Şekil 7.1. Veri madenciliği gelişim süreci	24
Şekil 7.2. Veri madenciliği süreçleri.....	25
Şekil 8.1. Örnek karar ağacı yapısı	31
Şekil 9.1. Weka açılış ekranı..	36
Şekil 9.2. Weka ön hazırlık sayfası	37
Şekil 9.3. Weka sınıflandırma ve sonuç ekranı.	41
Şekil 10.1. Kart limitinin gruplandırılmış gösterimi	46
Şekil 10.2. Maaş niteliğinin gruplandırılmış gösterimi	46
Şekil 10.3. Varlık miktarı niteliğinin gruplandırılmış Gösterimi	46
Şekil 10.4. Alışveriş miktarı ortalama niteliğinin gruplandırılmış gösterimi	46
Şekil 10.5. Alışveriş yiyecek niteliğinin kategorik gösterimi	47
Şekil 10.6. Alışveriş giyim niteliğinin kategorik gösterimi	47
Şekil 10.7. Alışveriş kozmetik niteliğinin kategorik gösterimi	47
Şekil 10.8. Alışveriş mücevher niteliğinin kategorik gösterimi	47
Şekil 10.9. Alışveriş eğitim niteliğinin kategorik gösterimi	48
Şekil 10.10. Alışveriş diğer niteliğinin kategorik gösterimi	48
Şekil 10.11. Girdi alışveriş çeşit niteliğinin kategorik gösterimi	48
Şekil 10.12. Girdi alışveriş miktar niteliğinin gruplandırılmış gösterimi	49
Şekil 10.13. Risk değerlerinin gösterimi	49
Şekil 10.14. Müşteri eğitim veri seti gösterimi	49
Şekil 10.15. Müşteri test veri seti gösterimi	50
Şekil 11.1. Weka karar ağacının yapısı	52
Şekil 11.2. Girdi alışveriş miktarı 100 – 1000 ve 1000 – 3000 kurallarının gösterimi	53
Şekil 11.3. Girdi alışveriş miktarı 3000 – 6000 kurallarının gösterimi	54

	Sayfa No.
Şekil 11.4. Girdi alışveriş miktarı 6000 – 10000 kurallarının gösterimi	55
Şekil 11.5. Girdi alışveriş miktarı 10000 – 20000 ve 20000 - 30000 kurallarının gösterimi	56
Şekil 11.6. Hata oranlarının gösterimi	57
Şekil 11.7 Risk değerlerine göre yapılacak işlemler	58
Şekil 12.1. Uygulama kullanıcı giriş ekranı.....	59
Şekil 12.2. Uygulama anasayfa ekranı... ..	60
Şekil 12.3. Uygulama sanal pos ekranı... ..	61
Şekil 12.4. Uygulama müşteri risk grubu ekranı.....	61
Şekil 12.5. Uygulama müşteri işlemleri ekranı.....	62
Şekil 12.6. Uygulama müşteri ekle ekranı	63
Şekil 12.7. Uygulama müşteri güncelle ekranı	63
Şekil 12.8. Uygulama müşteri sil ekranı	64
Şekil 12.9. Uygulama alışveriş listeleme ekranı	65

ÖZET

BANKACILIKTA DOLANDIRICILIK YÖNTEMLERİ ve TESPİTİ

Tez çalışmasında, temel olarak bankacılık ve finans sektöründe dolandırıcılık ve hırsızlığın önüne geçilmesi hedeflenmiştir. Bu amacı desteklemek için tez de bazı yazılım dilleri, veritabanı yönetim sistemi uygulamaları ve veri madenciliğinin ilgili model ve algoritmaları kullanılmıştır. Tez çalışmasındaki amaç hem maddi, manevi kayıplarının önüne geçmek hem de kurumlar için müşterilerine daha hızlı, güvenli hizmet sağlayarak şirketsel imajlarını korumaktır.

Tez çalışmasında temel olarak müşterilerin geçmiş dönemlerdeki alışveriş geçmişleri veritabanı sistemlerinde tutulmuştur. Bu alışverişler kategorize edilerek hangi müşteri parasını, hangi ürün yelpazesine harcadığı saptanarak veri tabanına kayıt edilmiştir. Kategorilerde ki ürünlere yatırdığı miktarlar da göz önünde bulundurulmuştur. Tez çalışmasında banka müşterilerinin beş aylık harcamaları kategori ve miktarlarla kayıt altına alınmaktadır. Müşteri kartı belirli yöntemlerle kopyalandığı ve çalındığı durumlarda yapacağı alışverişlerin alışkanlıklarına ters düştüğü durumlarda veri madenciliği ile oluşturulmuş ağaç sonucu oluşan yapı devreye girerek dolandırılma riski belirlenecektir.

Veri madenciliğinde sınıflandırma yönteminin içindeki karar ağacı ile sınıflandırma tekniğinin C4.5 algoritması kullanılmıştır. Karar ağacı tekniği ile gelişim eğitim verisi üzerinden sağlanmıştır.

Anahtar Kelimeler: Bankacılık, dolandırıcılık, sınıflandırma, karar ağacı

ABSTRACT

FRAUD METHODS IN BANKING AND ITS DETECTION

In this study, it has been aimed to prohibit fraud and theft in banking and finance sector basically. For supporting this purpose, some software languages, database management system applications and related model and algorithms of data mining have been used in this study. In the thesis study, main purpose is both prohibiting moral, material losses and protecting their corporate images by providing faster, secured service to clients for institutions.

In the thesis study, clients past shopping histories have been detained in database systems basically. These transactions have been saved in the database by categorizing and detecting which client spent his/her money to which product range? Imbursed amounts to products have been also considered in the categories. In the thesis study, bank clients five monthly expense are recorded with category and amounts. In case of copying and stealing the client's card with specific methods and contradiction his/her upcoming shopping to his/her behaviors; fraud risk will be designated by getting involved construct which is consisted because of tree that it is generated with data mining.

In the data mining, decision tree within classification method and classification technique's C4.5 algorithm have been used. Developing has been provided over training data with decision tree technique.

Keywords: Banking, fraud, classification, decision tree

1. GİRİŞ

Gelişmekte olan küresel dünyada insanların belirli ihtiyaçları zamanla önemini arttırmıştır. Artık insanlar için zaman, hız ve güvenlik gibi kavramlar hayatın olmaz ise olmazlarından. Bankacılık ve finans kavramlarında bu nedenler ön plana alınarak değerini ve kapsamını sürekli arttırmıştır. Para kavramının bu denli önemi para eşittir banka formülünü akıllara getirmiştir. İki yapının gelişmesi global dünyada doğru orantılıdır. İnternet sisteminin gelişmesi ve sanal dünyanın insanlığı sarmasıyla bankalarda internet bankacılığı adı altında sanal dünyaya adımlarını atmıştır. Online bankacılığa geçen bankalar bu sanal yapı sayesinde zaten üstüne çektikleri dolandırıcılık tehditlerini misliyle arttırmıştır. Bu gelişme ile artık bankalar müşterilerine daha hızlı, kaliteli ve kullanıcı dostu sistemler sağlamakta yani hayatı kolaylaştırmakla beraber belirli güvenlik tehditlerini de yanında getirmiştir. Güvenlik tehditleri genel olarak;

- i. Kredi kartı bilgilerinin ele geçirilmesi.
- ii. Çağrı merkezleri yoluyla dolandırıcılık.
- iii. Ödeme sistemleri dolandırıcılığı

gibi birçok tehdit artık bankalar için geçerli ve önemlidir.

Bankalar, gelen bu tehditleri yok etmek veya önleyebilmek için yine aynı tehditleri oluşturan teknolojik gelişmeleri kullanmak zorunda kalmıştır. Bunlar;

- i. Mail uygulamaları
- ii. Şifreler
- iii. Güvenlik soruları
- iv. Yapay zeka algoritmaları
- v. Anlık SMS

Ancak ne yapılırsa yapılsın ne kadar dolandırıcılık yöntemleri tespit edilirse edilsin ve ne kadar önleme uygulamaları olursa olsun artık bankalar karşılıklı ve sürekli gelişim evresine girmişlerdir. Bu evrede dolandırıcılar yeni tehditler yaratma, bankalar yeni önlemler almak zorundadır.

Tez de dolandırıcılığın direkt tespiti ve buna bağlı olarak önlenmesi gibi işlemler gerçekleştirilmiştir. Tez kapsamında dolandırıcılık öncesine yönelik temel bir önlem algoritması bulunmamaktadır. Müşterinin kredi kartının kopyalanması ve bilgilerinin ele geçirilmesi kısmında geliştirilecek algoritma etkin değildir. Tez dolandırıcının kredi kartı üzerinde işlem yapacağı sırada aktifleşecektir. Bu işlemler şunlar olabilir;

- i. Yeni alışveriş işlemlerinin gerçekleştirilmesi
- ii. Banka hesaplarından para çekimi

Tez kapsamında müşterinin belirli zaman aralıkları arasındaki işlemleri (altı aylık, bir yıllık, beş yıllık) gibi periyotlar halinde oluşturulup sistemin veritabanında tutulması gerçekleştirildi. Müşterinin alışveriş geçmişi belirli kategorilere ayrıldı. Bu kategoriler;

- i. Giyim
- ii. Mücevherat
- iii. Diğer
- iv. Eğitim
- v. Market
- vi. Kozmetik

Oluşturulan kategoriler müşteri alışverişi ile ilişkilendirilmiştir. İlişkilendirme sadece eşleştirme şeklinde olmayıp yapılan alışveriş miktarında sisteme eklenerek müşterinin alışkanlıkları belirlenmiştir. Alışverişlerde konum bazlılık kullanılabilir. Yani alışveriş yapılan konumlarda sisteme eklenebilir.

Bilgiler sistemde kalacak şekilde kredi kartı kopyalanan veya çalınan müşterinin, dolandırıcı veya hırsız tarafından yapılacak alışverişleri sisteme eklenerek sistem aykırı durumlar var ise saptanacak ve buna karşı olarak müşterinin risk grubuyla ilgili bilgilendirme sağlanacaktır. Belirlenen risk grubu üzerinde önleme çalışmaları da yapılabilir. Bu önleme işlemleri şunlar olabilir;

- i. Müşteri telefonla SMS gönderimi.
- ii. E-posta ile bilgilendirme, onay kodu.
- iii. Hesabın güvenliğinden emin olana kadar hesabı askıya alma.

Tez çalışmasında başka bir ayrıntı daha bulunmaktadır. Bu ayrıntı müşteri kendi alışveriş geçmişine aykırı işlemleri kendisinde gerçekleştirebilir. Bunun için müşterinin alışveriş alışkanlıklarının belirlenmiştir. Belirleme işlemi veri madenciliğinin uygun model ve algoritmaları ile gerçekleştirildi. Hata değerlerini ve kritellerini belirlenmiş gerekli işlem yapılarak müşteri mağduriyetinin önüne geçilmiştir. Gerekli işlemler yapılmaz ise daha güvenli sistem olsa dahi hız ve zaman gibi konularda kullanıcıyı tatmin etmiyerek bankayla müşteri arasında belirli problemlerinin ortaya çıkmasına olanak sağlayabilir.

Geliştirilen tezin kapmasına gelinirse tez çalışmasında kullanıcı kendi alışveriş işlemini gerçekleştirmesi için bir arayüz gerekmektedir. Gereken arayüz belirli programlama teknikleri ile oluşturuldu. Oluşturulan arayüzün arkasına veri madenciliği kullanarak elde edilen karar ağacı modeli eklendi. Sistem, gelişmeye ve büyümeye doğrudan açık olmalıdır. Modellenmiş yapıda örnek seneryolar ile sisteme karar verme ve algılama yetenekleri kazandırıldı.

Literatürün ikinci bölümünde bankacılığın tanımı, gelişimi, ülkemiz ve dünyada ki tarihçesi anlatılmıştır. Üçüncü kısımda ki internet bankacılığı bölümünde bu tip bankacılığın işlemleri, avantaj ve dezavantaj bilgileri tanımlanarak açıklandı. Bununla beraber dördüncü kısımda sosyal mühendislik kavramıyla bankacılıkta ki dolandırıcılık kavramı ifade edildi. Beşinci bölümde ise dolandırıcılık ve türleri konularından bahsedim sağlanarak altıncı bölüm ile bu dolandırıcılık yöntemleri ile dolandırıcılık tespiti nasıl yapılır kısımlarının aktarımı sağlandı. Yedinci kısımda veri madenciliği başlığı altında tanımı ve keşif süreçlerinden başlayarak tarihçe, model ve yöntemler aktarıldı ve devam yolu olarak sekizinci kısım ile sınıflandırma yöntemi ve buna uygun algoritmaların anlatımı sağlanmıştır.

Dokuzuncu kısımda ki gereç ve yöntem bölümünde tez kapsamında kullanılan teknoloji ve diller ne olduğu ayrıca kullanılma sebeplerinden bahsedildi. Onuncu kısım üzerinde ki veri setinin hazırlanması bölümünde veri setinin nerden elde edildiği ve nitelik ve sınıf değerleri ayrıntılı olarak açıklanmıştır. Onbirinci kısım ile işleme sokulan veri setine dair sonuçlar, oluşan kurallar ve hata oranları bulgular başlığı

altında sunuldu. On ikinci bölüm üzerinde tez çalışması üzerinde gerçekleştirimi sağlanan arayüzün görsellerinin sunumu yapılmıştır.

Onüçüncü bölüm ile geçmiş çalışmalarla yapılmış tez karşılaştırılarak avantajlarından tartışma başlığı altında bahsedilmektedir. Son bölümde ise tez yapımı sonucu elde edilen sonuç değerleri ile beraber nasıl bir fayda sağladığı sonuç başlığı altında ifade edildi.



2. DÜNYADA ve TÜRKİYE'DEKİ BANKACILIK GELİŞİMİ

Tüm dünyada bankacılık kelimesi İtalyanca “banco” kelimesinden gelmektedir. Fakat diğer tüm dillerde bu kelime ufak oynamalarla küçük farklılaşmalar göstermektedir. Lambordiyalı yahudiler bankacılığa ilk atımlarına pazarlarda kurdukları beyaz masalarla başlamışlardır. Finansal sistem temel olarak iki parçadan oluşmaktadır. Bunlar;

- i. Tahvil ve bono piyasaları
- ii. Finansal araçlar (Sigorta şirketleri, sosyal güvenlik kurumları)

Bankacılıktaki gelişim temel olarak para kavramıyla doğrudan bağlantılıdır. Başka bir yolla bu ilişki ifade edilmek istenilirse para eşittir banka formülü ile piyasada ne zaman para akışı fazlaşır ise bankacılıktaki gelişim doğrudan etkilenir.

Bankacılık hareketlerinin ilk kilometre taşı tahmini olarak M.Ö. 3500 yıllarında başlamaktadır. Mezopotamya'daki “Kızıl Tapınak” (M.Ö.3400-3200) tarihteki ilk banka olarak bilinmektedir.

Türkiye 'de ilk bankacılık faaliyetleri ve ilk banka yapısı Osmanlı İmparatorluğu dönemine dayanmaktadır. Osmanlı İmparatorluğu döneminde 1847 yılında İstanbul bankası adı ile Galata Bankerleri tarafından ilk Türk bankasının kuruluş adımları atılmıştır. İlerleyen süre zarfında 1852 'de bankacılık faaliyetlerine son vermek ile beraber kapatılmıştır. Osmanlı İmparatorluğunun yıkılması ve Türkiye Cumhuriyetinin kurulması ile beraber, Cumhuriyet döneminin ilk doğduğu zamanlarda yirmi iki tanesi milli, on üç tanesi yurtdışı kaynaklı otuz beş banka bulunmaktaydı (<https://www.tarihnotlari.com.>, Erişim Tarihi: 11 Şubat 2019).

3. İNTERNET BANKACILIĞI

Dünyadaki gelişim ve büyüme bir çok sektörün internet ve sanal ortamlara buluşmasını sağladı. İnternet teknolojisine geçiş firmaları bir takım avantajlar sağlamıştır. Bu avantajlar aşağıdakilerdir;

- i. Hız
- ii. Daha fazla müşteri kapasitesi
- iii. İşçi sayısında azalmalar
- iv. Ekonomik kazançlar
- v. 7/24 Bankacılık işlemleri

Yukarıda belirtilen kavramlar düşünüldüğünde bankacılık ve finans sektöründe bu gelişmelere ayak uydurarak internet dünyasına giriş yapmıştır. İnternet bankacılığı dünyada 1983 yılında “The Bank of Scotland” bankası ile başlamıştır.

İnternet bankacılığı kavramı, bankalarda yapılan bir çok iş için müşteri hesaplarını uygun hale getirerek insanların elektronik ortama geçiş yapmasını sağlamıştır. Müşteriler ise ellerinde bulunan bir takım bilgiler ve şifreler ile kolayca bankacılık işlemlerini gerçekleştirebilmektedir. Bu işlemler ilk başta sadece PC 'lerde gerçekleştirilebilirken artık telefonlar, tabletler vb. birçok cihaz ile gerçekleştirebilmektedir (<https://www.pcnet.com.tr>., Erişim Tarihi: 16 Ocak 2019).

3.1. İnternet Bankacılığı İşlemleri

İnternet bankacılığının yardımı ile günümüzde bir takım işlemlerin internet üzerinden yapılması mümkün olmuştur. Bu şekilde işlem zamanının kısaldığı doğrudur. Yani internet bankacılığı kavramı ile bankalardaki yoğunluk azalmıştır. On

beş yılı aşkın süredir kullanımı devam etmekte olup kullanımı günbegün gelişmektedir.

İnternet bankacılığı ile hangi hizmetler alınmaktadır;

- i. **EFT / Havale:** Belirlenen kişi ile internet üzerinden hesap yada IBAN numarası ile para aktarma işlemidir.
- ii. **Hesapların Kontrolü:** İnsanların banka üzerinde tüm hesaplarını anında görüntüleyebilme ve toplu kontrol gibi imkanlarını müşterilere kazandırır.
- iii. **Ödeme İşlemler:** Aylık ciddi zaman kayıplarına uğranılan fatura ödemeleri gibi işlemlerdir.
- iv. **Yatırım İşlemleri:** Hesaplar üzerinde gerekli yatırım işlemlerini yapılan yerdir.

3.2. İnternet Bankacılığı Avantaj ve Dezavantajları

Bankacılık ve finans sektörünün internet bankacılığına geçişi bir takım avantaj ve dezavantaj sağlamaktadır. İnternet bankacılığının avantajları aşağıda belirtildiği gibidir;

- i. **Zaman Kazancı:** İnternet bankacılığının sağladığı en büyük yarardır. Artık bankalarda sıra bekleme ve işlem esnasında ki zaman kayıpları azalmıştır.
- ii. **Masrafların Düşürülmesi:** Bankacılık işlemlerini gerçekleştirmek için bankalara belirli miktarda ekstra para verilir ama kurumlar internet bankacılığının kullanımının arttırılması için, internetten yapılan işlemlerinin ücretlerini düşürmüştür.
- iii. **Zaman Bağımsız İşlem Yapabilme Özelliği:** Artık müşteriler işlemlerini bankalara gitmeden zamandan bağımsız olarak gerçekleştirebilmektedir.
- iv. **Kolay Yapı:** İnternet bankacılığı sistemlerinde, insanlara kullanıcı dostu arayüzler sunularak daha kolay işlem yapabilme imkanı sağlar.
- v. **Hızlı Kredi Alma İmkanları:** Banka müşterisinin kredi geçmişi, kredi sicili bununla beraber maaş bordrosu göz önünde bulundurularak sistem tarafından müşteriye kredi imkanının sağlanması olayıdır.

İnternet bankacılığının dezavantajları aşağıda belirtildiği gibidir;

- i. **Güvenlik:** Çevrimiçi bankacılığın en büyük tehdidi ve dezavantajı güvenlik zaafiyetidir. Müşterinin özlük bilgileri ve birikimlerinin kötü niyetli insanların emelleri altındadır.
- ii. **Müşteri Hizmetleri:** Geleneksel bir müşteri ise işlemlerini bankalarda konuşarak halletmektedir. Bu yöntem ile sorunlarına anında cevap alır. Fakat insanların hayatlarına giren müşteri hizmetleri kavramıyla zaman kayıplarının yaşanması olasıdır.
- iii. **Erişilebilirlik:** Bankacılık işlemlerinin yapılabilmesi için internetin her zaman açık olması gerekir. Anlık kopma ve ağdaki yavaşlıklar işlemlerin hızını doğrudan etkiler.
- iv. **ATM Bakımları:** Kullanılan bankanın ATM lerinde sıkıntı var ise başka banka ATM lerini kullanılır. Kullanım için ekstra ücret talep edebilirler buda banka müşterileri için bir dezavantajdır (Bhardwaj, Gupta, 2016).

4. SOSYAL MÜHENDİSLİK

İngilizce tanımı ile “Social Engineering” en temel dolandırıcılık yöntemlerinden biridir. Sosyal mühendislik yapısı ilk olarak insan ilişki ve dikkatsizliklerini kollayarak kullanıcıların hakkında bilgi ve belge toplayarak çeşitli yöntemlerle bunları kullanmaya dayanır. Buradaki amaçları saldırı esnasında kendilerine yardımcı dokunacak yönetici, çalışan ve müşterilerin kişisel bilgilerini, şifreler ve şirketlerin yapısı hakkında ki her türlü bilginin toplanmasıdır (<https://www.tbb.org.tr.>, Erişim Tarihi: 22 Ocak 2019).

Bütün sosyal mühendislik yöntemlerinde insan hayatı, önyargılar üzerine dayandırılır. Önyargılar insanları karar verme konusunda etkileyerek dolandırıcıya bir takım avantajlar sağlar. İnsanlarda oluşması muhtemel bir kaç önyargı aşağıda belirtilmiştir;

- i. Otorite Korkusu
- ii. Kaybetme Korkusu
- iii. Belirli Sözler
- iv. Kazanma Hissi
- v. Değerli Olma Hissi

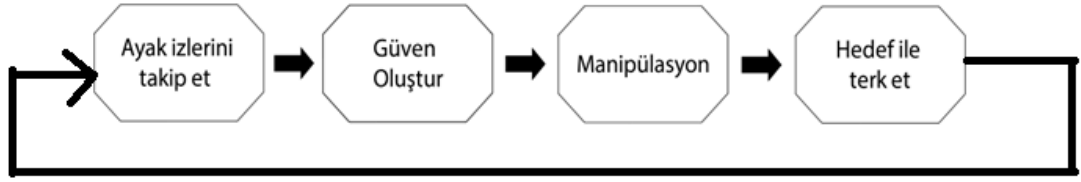
Sosyal mühendislik “hack işlemlerinin atası olarak kabul edilir.” Bu kavramın mucidi Kevin Mitnick sosyal mühendislik kavramını şu kelime ile açıklamıştır.

“Bir kurumun güvenliği için en büyük tehlike anahtar bir programdaki güncelleme eksikliği ya da kötü yüklenmiş güvenlik uygulaması değildir. Gerçekte en büyük tehlike siz olabilirsiniz. Şahsen, insanı yanıltmanın teknolojiyi yanıltmaktan

daha kolay olduğunu gördüm. Organizasyonların bir çoğu insan faktörünü göz ardı etmektedir” (<https://www.tbb.org.tr>., Erişim Tarihi: 22 Ocak 2019).

4.1. Yaşam Döngüsü

Sosyal mühendisler müşteri kimlik bilgileri, şifreler vb. birçok bilgiyi ele geçirmeyi planlar. Bunları belirli bir döngü içerisinde yapar. Buna yaşam döngüsü denir. Sosyal mühendislik yaşam döngüsü Şekil 4.1. 'de gösterilmiştir.



Şekil 4.1. Sosyal mühendislik yaşam döngüsü

- i. **Ayak İzlerini Takip Et:** Sosyal mühendis bilgilerini, şifrelerini ele geçirmek istediği müşterinin alışkanlıklarını belirliyerek onunla ilgili ilk etapta öğrenebileceklerini belirler.
- ii. **Güven Oluşturma:** Önceden planlanmış olay kurgusuna göre kişilerle ilişkiye geçilir güven ortamı sağlanacak ortam hazırlanır. Örneğin; ilgili kişi kendini kurum çalışanı olarak tanıtmayı ile beraber karşı taraftan güven kazanmaya çalışır.
- iii. **Manipülasyon:** Olay kurgusunu baz alarak istediği hedefe ulaşmak için karşı taraftan aldığı güveni manipüle ederek bilgi elde etmeye çalışmasına verilen isimdir.
- iv. **Hedef İle Ortamı Terk Etmek:** Olay kurgusunda hedeflenen bilgilerin ele geçirilmesi için çalışılarak ve bu olay kurgusunun sonunda herhangi bir şüphe uyandırmaksızın ortamdaki geri çekinilebilmesidir.

4.2. Kullanılan Genel Yöntemler

Sosyal mühendislik kullanılan yöntemlere göre sınıflandırıldığında iki alt başlık halinde incelenebilir;

4.2.1. Temeli Teknoloji Olan Sosyal Mühendislik

Temeli teknoloji olan sosyal mühendislikte bilgisayar üzerinden kişiye ait olan bilgiler elde edilebilir. Bu bilgiler müşterinin özlük bilgileri, şifreleri vb. olabilir.

4.2.2. Temeli İnsan Olan Sosyal Mühendislik

Bu yöntemde bilgiler teknoloji yoluyla değil doğrudan müşterinin kendisi tarafından sosyal mühendislik yöntemleri ile elde edilir. Burada taklit, ikna etme vb. metotlar uygulanır. Bazı yöntemler, aşağıdaki maddeler ile açıklanmıştır.

Bu maddeler;

- i. Müşterinin Taklit Edilmesi:** Bu yapıda sosyal mühendis bilgilerini elde edeceği kişinin algısını yanıltmaya çalışır.
- ii. Teknisyen Yöntemi:** Sosyal mühendis kendini bir şirketin elemanı göstererek dolandırıcılık kişi veya kurumun bilgilerini öğrenmesine verilen isimdir.
- iii. Otorite Korkusu:** Sosyal mühendis kendini devlet görevlisi (Savcı, polis, hakim) gibi tanıtarak insanların korkularından yararlanmasına verilen isimdir.
- iv. Kaybetme Korkusu:** Burada sosyal mühendis dolandıracağı kişiye kısa vadeli fırsat sunma yolu ile müşteriye bilgilerini vermeye zorlamasıdır.
- v. Kazanma Hissi:** Bu yapıda internet siteleri, kısa mesajlar yardımıyla kişilerin bir şeyler kazandırdığına ikna ederek bilgilerini almaya verilen isimdir.
- vi. Ters Psikoloji:** Sosyal mühendis kendini dolandırılacak kişiye verdiği yardımlar ile güvenini sağlamaya çalışır. Bu yöntem bilimle kişinin özlük bilgilerini ve şifre yapılığını öğrenmesidir.
- vii. Elektronik Posta Dolandırıcılığı:** Günümüzde popüler olarak kullanılan bir yöntemdir. Bu yöntemle müşteriye belirli imkanlar sunarak ön ödeme yapması sağlanır.

4.3. Őirket Őzeri Etkiler

Bu mŕhendislik yapısı Őirketler ũzerinde bazı olumsuz kalıcı etkiler oluŐturabilir.

- i. İmaj Zedelenmesi
- ii. Maddi Zararlar
- iii. Hukuki Cezalar
- iv. Vakit Kaybı
- v. Artan Gŕvenlik Maliyetleri



5. DOLANDIRICILIK VE TÜRLERİ

Temel anlamıyla dolandırıcılık kavramı, kandırma işlemi üzerinden yapılan eyleme verilen isimdir. Bu eylemi gerçekleştiren kişiye de dolandırıcı adı verilir. Yapılan eylem sonucunda dolandırıcı kendi veya başka insanlar üzerine maddi kaynak aktarımı sağlayabilir. Türk ceza kanununun 157. Maddesine göre de dolandırıcılık eylemini gerçekleştiren kişilere bir yıldan beş yıla kadar hapis cezası ve beş bin gün kadar adli para cezası verilir. Günümüzde internet kullanımının sürekli artmasıyla, dolandırıcılar sanal ortama geçerek bankacılık sektörü üzerinden yapılan dolandırıcılık işlem sayısını katbekat artırmıştır. Bankacılık dünyasında en fazla yapılan dolandırıcılık Türleri aşağıdaki gibidir;

- i. İnternet Dolandırıcılığı
- ii. Çağrı Merkezli Dolandırıcılıklar
- iii. Kartlı Sistem Dolandırıcılığı

5.1. İnternet Dolandırıcılığı

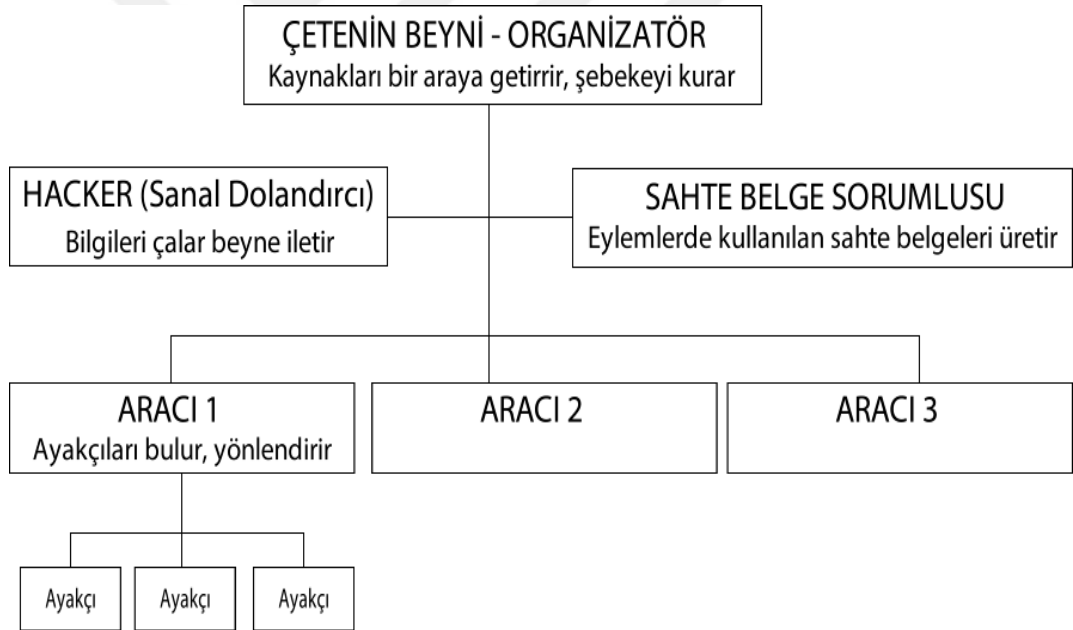
Bilgisayar sistemleri üzerinde kendisine ait olmayan gizli bilgileri öğrenmek, bunlar üzerinde işlemler yapmak kişilere, kurum veya kuruluşlara zarar vermek amacıyla işlem ve çalışma yapan bilgi hırsızlarına “hacker” denir. Hackerlerin temel hedefi hakkı olmadan elde ettiği bilgileri kullanarak bir takım faydalar sağlamaktır.

Günümüzde sanal alamin gelişimiyle bilgi hırsızlarının faaliyetleri artmıştır. Birçok etkili yöntemle kişi kimlik bilgileri, şifreler, telefon numaraları vb. verileri elde edebilirler hatta sadece öğrenmek ile kalmayıp PC ve serverlara kalıcı zararlar verebilirler. Bilgi hırsızları çoğu zaman kişilerin internet hakkında ki bilgi eksikleri ve alışkanlıklarından faydalanırlar. İnternet üzerinden yapılan dolandırıcılıkta

kullanıcının bilgisayarına virus, trojen vb. gibi yazılımlar sokmalıdır. Bunların kişilerin bilgisayarına dosyalar, mailer, uygulamalar, müzikler vb. bizzürü yolla yükleyebilirler. Unutulmamalıdır ki hiçbir dolandırıcılık programı bir PC 'ye kendiliğinden kurulamaz.

Yukardaki bilgiler ışığında bilgi hırsızları için en etkili ve kazançlı yerler banka ve finans sektörüdür. Günümüz Türkiyesin'de hemen hemen bütün bankalar internet bankacılığı hizmeti vermektedir. Hizmeti veren bankalar için en büyük tehdit bilgi hırsızlarıdır. Tehditleri farkederek bankalar güvenlik konusuna büyük yatırımlar yapmıştır. Yatırımlar sorunları azaltsada tam olarak çözüme kavuşturmak imkansızdır (Özbay, 2007)

İnternet dolandırıcılığı tek bir kişi tarafından kolayca yapılabilecek bir eylem değildir. Kötü niyeti insanlar ekip halinde çalışmaktadır. İnternet dolandırıcılığı yapan hırsızları ifade eden örnek şema Şekil 5.1. 'deki gibidir.



Şekil 5.1. Dolandırıcılık yapan çetenin organizasyonu

5.1.1. Bilgilerin İnternet Üzerinden Düzenlenmesi

Bilgi hırsızların kullanıcıların özlük bilgileri, şifreleri vb. birçok bilgileri bir takım yollarla elde edebilirler. İnternet üzerinden elde ettikleri bazı yöntemler aşağıda belirtilmiştir;

i. Keylogger ve Screenlogger: Keylogger yapısı temel olarak yüklendiği bilgisayarda klavye hareketlerini takip ederek, kayıt altına alır. Kayıt altına alınan

klavye yazıları bilgi hırsızına e-mail gibi yollarla aktarılır. Kayıt altına alınan bilgiler ‘.txt’ uzantılı metin dosyaları şeklinde tutulmaktadır. Screenlogger yapısı adındanda anlaşıldığı üzere klavye yazılarını değil ekran görüntülerini alır. Aynı şekilde hırsıza e-posta yoluyla aktarım gerçekleştirilir. İki yapı birbiri ile eş güdümlü çalışır. Bu tip uygulamalara örnek olarak: Perfect keylogger, Phantom vb. verilebilir.

ii. Trojan Yazılımlar: Bu uygulama türkçe truva atı olarak adlandırılır. Truva yazılımları bilgisayar sisteminin çalışmasını sağlayarak uygulamayı bozamaz. Aynı şekilde şifre ve kullanıcı bilgilerinin elde edilmesi konusunda kullanılır. Bilgisayarlara bulaşımı oyun, uygulama, metin dosyası vb. gibi dosyaların indirilmesi ile bilgisayara aktarımı sağlanır. Bu tip yazılımlara örnek olarak: Netbus, Subseven ve Prorat verilebilir.

iii. Pop-Up Ekran Dolandırıcılığı: Bu yazılım tipinde kullanıcı internet üzerinde gezinirken, sayfalarda ekstra ekranlar çıkar buradaki ekranları kullanıcı kullanarak PC’ye dosya indirmesi vb. yollarla kullanıcı bilgileri çalınabilir.

iv. Spam Mailler: Bu yapı çoğunlukla maillerimize bakarken yanlarında gözükten ilk başta zararsız sandığımız bu postalar içlerinde zararlı yazılımlar (Keylogger, Screenlogger, Trojan, Virüs) vb. bulundurabilir.

v. Donanımlar: Donanımlar bilgisayara takılan parçalar, dinleyici veya kaydedici gibi fiziksel parçalar olabilir. Bunlar kurum içerisinde çalışan insanlar veya kuruma dışardan gelen insanlar tarafından etrafa yerleştirilebilir. Man in The Browser Attack (MitB); bu yöntemde temel amaç kullanıcı bilgilerini almaktan ziyade tarayıcı üstünde saldırı yapmaktır. Bu yazılımlara örnek vericek olursak zeus ve snowal gibi programlardır (<https://www.tbb.org.tr>., Erişim Tarihi: 22 Ocak 2019).

5.1.2. İnternet Güvenliği İçin Neler Yapılmalı

Bankacılıkta internet üzerinde işlem yapan kullanıcılar belirlenen işlemleri bazı zaman aralıklarında yaparak kendi güvenliklerini arttırabilirler. Bu işlemlerden birkaçı aşağıda belirtilmiştir;

i. Özlük Bilgileri: Bunlar kullanıcı adı, parola, doğrulama kodları vb. gibi bilgilerdir.

ii. Kredi Kartı Malumatı: Bunlar kart numaraları, son kullanma tarihi ve cv gibi bilgilerdir. Şifreler karmaşık olmalıdır. Kullanıcı kendi ile direkt alakalı basit şifreler kullanmamaya dikkat etmelidir.

iii. Şifreler: Kullanıcı, şifreleri kullandığı her platformda aynı olmamalıdır. Olması durumunda tek şifresi ele geçirildiğinde bütün hesapları deşifre olmaktadır.

iv. Kullanılan Ağlar: Kullanıcımız sadece bildiği ve güvendiği bilgisayar ve ağlarda yapacağı işlemleri gerçekleştirmelidir. Örneğin; internet kafe veya ücretsiz ortak ağların kullandığı yerlerde önemli işlemler yapılmamalıdır.

v. Lisanslı Yazılımlar: Crackli yazılım olarak isimlendirilen programların içine casus programlar yüklenmiş olabilir. Bunlar düşünülerek lisanslı yazılımlar kullanılmalıdır.

Anti-virüs Yazılımları: Bilgisayarlarımızda anti-virüs yazılımları kullanılmalıdır. Bu yazılımlar casus programları önlemede oldukça başarılıdır (http://adudspace.adu.edu.tr., Erişim Tarihi: 01 Şubat 2019).

5.2. Çağrı Merkezli Dolandırıcılıklar

Kurumların iletişim halinde buldukları kişi veya kurumlarla olan ilişkilerinin düzenlendiği yere çağrı merkezleri denir. Bu merkezler müşterilere 7 / 24 hizmet sağlayarak sorunlarını çözmekte etkin olarak rol oynarlar. Çağrı merkezlerinde müşterilere verilen bazı hizmetler aşağıdaki gibidir;

- i. Bankacılık
 - a. Fatura Ödeme
 - b. Havale, EFT
 - c. Otomatik Talimatlar
- ii. Kredi Kartı
 - a. Ekstre borcu, güncel hesap işlemleri
 - b. Kimlik bilgilerinin güncellenmesi
 - c. Kart iptali
 - d. Borç ödeme işlemleri
 - e. Şifre İşlemleri

5.2.1. Dolandırıcı Tarafından Çağrı Merkezinin Aranması

Bu dolandırıcılık yönteminde bilgi hırsızı çalışanın zaafiyetlerini veya sosyal mühendislik teknikleri kullanabilir. Müşteri temsilcileri ilgili abone veya o kurum ile çalışan müşterilerin belirli soru dizinlerini kullanarak müşteriye cevap vermektedirler. İşlem yapılması gereken kritik bilgileri müşteriye sorularak istenir ve sonrasında dolandırıcılar zamanla müşteri temsilcilerinin davranışlarını taklit etmesiyle normalde şubeden yada çağrı merkezi tarafından alınabilecek bilgileri müşteriden isteyerek dolandırmaktadırlar.

Çağrı merkezinin güvenlik aşamaları vardır. Bu aşamalara örnek vermek gerekirse annenizin kızlık soyadı veya nüfus cüzdanı seri no gibi kritik güvenlik soruları sorularak işlemler sağlanmaktadır. Burada bir hata gerçekleşirse bunun ötesinde işlem yapılması için yetki verilmez. Bazı durumlarda da müşterinin kişisel elektronik adresi ya da telefonuna bilgilendirme, doğrulama bilgileri gönderilir. Bunun dışında müşteriye sorulan bilgiler kısıtlı kalmaktadır (<https://www.tbb.org.tr>, Erişim Tarihi: 22 Ocak 2019).

5.3. Kartlı Sistem Dolandırıcılığı

Banka kullanıcılarının ATM ve POS cihazlarını kullanmaya başlaması, 1980-1990 yılları arasında başlamıştır ve dünyanın globalleşmesi ile kullanımı hızla artmıştır. İlk yıllarda dolandırıcılar daha ilkel yöntemleri kullanmıştır. Fakat teknolojinin gelişimi ve müşterilerin bilgilenebilmesiyle daha gelişmiş yöntemler kullanmak zorunda kalmışlardır (<https://www.tbb.org.tr>, Erişim Tarihi: 22 Ocak 2019).

5.3.1. ATM Dolandırıcılığı

Atm sistemleri üzerinden dolandırıcılık işlemi yapılırken aşağıdaki yöntemler kullanılabilir. Bu yöntemler;

i. Kart Sıkıştırma: ATM lerde gerekli işlemler yapılır iken müşterinin kredi kartı makinenin içinde kalır. Dolandırıcılar buralara kağıt, yapışkan madde vb. ile kartı ATM içerisinde bırakabilirler.

ii. Zararlı Programlar Yükleme: Bilgi hırsızları teknolojinin gelişimi ile ATM üzerindeki bilgisayarın içine yükledikleri zararlı yazılımlarla makine içindeki paraları çekebilirler. Örnek olarak “Tyupkin” yazılımı verilebilir.

5.3.2. Kart Dolandırıcılığı

Kartlı sistemlerde yapılabilecek dolandırıcılık yöntemleri;

i. Account Takeover: Bu yöntemde temel olarak hesaba bağlı bakiyenin boşaltılması amaçlanmıştır. Yöntem bilim olarak müşterinin temel bilgilerinin bilinmesi yeterlidir.

ii. Çalıntı: Burada fiziksel olarak hırsız kredi kartı sahibinin kartını çalarak kaçar. Kullanılan en basit yöntemdir.

iii. Kart Kopyalaması: Bilgi hırsızları ATM lerde yerleştirdikleri özelleşmiş cihazlar ile müşterinin kredi kartını kopyalayabilirler. Bu cihazlara “skimmer” adı verilir. Temel olarak bu parça sadece kartı kopyalar. Şifreyi öğrenmek için PINPAD veya klavyeyi gören kameralar kullanılabilir.

5.3.3. Kart Dolandırıcılığından Korunma

Bu dolandırıcılık tipinde yapılan işlemler genel olarak ATM lerde gerçekleştirilir. Bunun önüne geçmek için aşağıdaki önlemleri alınabilir;

- i. Özlük bilgilerimiz kesinlikle saklı tutulmalıdır.
- ii. Banka ve finans sektörü dışında kalan bütün platformlar dikkatli kullanılarak, benzer şifreler kullanılmamalıdır.
- iii. Güvenliğinden tereddüt edilen ağlar ve PC lerde kesinlikle işlem yapılmamalıdır.
- iv. Bilinmeyen firmalar ile alışveriş yapılmamalı ve kişisel bilgiler onlarla paylaşılmamalıdır.
- v. İletişim bilgilerinin gizliliği sağlanmalıdır.

6. BANKACILIKTA DOLANDIRICILIK TESPİT YÖNTEMLERİ

Teknoloji ve internet dünyasının sürekli ve doğrudan gelişimiyle bankacılık sektörü de bu dijital ortama adımlarını atmıştır. Bankacılığın internet ortamına geçişiyle dolandırıcılık tehditleri misliyle arttı. Oluşan kayıpların önüne geçmek için bankalar gerekli güvenlik önlemlerini almıştır. Dolandırıcıların tespit edilmesi için bazı yöntemler aşağıda belirtilmiştir;

- i. İnternet üzerinden yapılan alışverişlerde müşteri IP numarasının tutulması ile dolandırıcı tespit edilebilir. (IP numarası, servis sağlayıcısından alınabilir.) Bu yapı hem önleyici model olarak hem de eylem sonucu oluşan dolandırıcılığın tespitinde kullanılır.
- ii. Bir bankadan yapılan para transferi gibi işlemlerde bankadaki güvenlik sistemleri ve para transferini yapan kişinin kimliğiyle dolandırıcılık tespit edilebilir.
- iii. Para transferi, havale gibi işlemlerde ATM üzerinden dolandırıcılık yapılır ise sistem kameralarından hırsız tespit edilebilir. Bu yapı fiziksel bir yapıdır. Önceden kestirim gibi önleyici bir işlem yapamamaktadır. Dolandırıcının yaptığı para transferi ve havale gibi işlemler sonucunda müşterinin dolandırıldığının sonradan farkına varması gibi durumlarda dolandırıldığını ispat etme gibi durumlarda kullanılır.
- iv. Para transferi, havale gibi işlemlerde internet bankacılığı üzerinden dolandırıcılık yapılır ise müşteri alışkanlıkları baz alınarak ortalama para transfer miktarı, yeni kişilere gönderme sıklığı ve gönderdiği kişilere olan aylık gönderim sayısı ve bunun gibi birçok değer tespit edilerek dolandırıcılığın önceden belirlenmesi ile dolandırıcılık eyleminin önüne geçilebilir.

- v. Alışveriş banka dışında fiziksel bir mağazadan yapıldıysa işletmenin barındırdığı güvenlik önlemleri ile dolandırıcılık tespit edilebilir. Güvenlik önlemleri kameralar üzerine kurulu olabilir veya şirket bünyesindeki bazı yazılımlar bu işlemi sağlayabilir.
- vi. Müşteri ve dolandırıcının hırsızlık anındaki yapabileceği telefon görüşmeleri ile dolandırıcılık tespit edilebilir. Burada şebeke operatörlerin yardımı oldukça önemlidir. Bu yapı daha çok fiziksel bir yöntemdir alışveriş işlemi tamamlandıktan sonra yapılan bir takım araştırmalar sonucu yukarıda belirtilen şebeke operatörleri yardımıyla dolandırıcılık tespit edilebilir.
- vii. Müşterinin genelde tercih ettiği mağazalar ve ikamet ettiği yerin dışında ani farklılıkla yaptığı alışverişler ile hırsızlığa uğradığı tespit edilir Bu tespitin yapılabilmesini, insanların genel olarak yaşadığı bölge üzerinde aynı lokasyonları tercih etmesi gibi nedenler sağlamaktadır. Genel olarak insan yapısı aynı lokasyonlar üzerinde alışveriş eylemini gerçekleştirme yönündedir. Ayrıca konum bazlı çalışarak son alışverişini yaptığı yer ile bir sonraki alışverişini arasındaki konum arasında mantıklı bir yakınlık yok ise dolandırıcılığı bu yöntemle tespit edilmesi de olasıdır (Şahin, Duman, 2011).
- viii. Müşteri alışkanlıkları dışında bir alışveriş veya aylık alışveriş miktarını aşırı geçen satın alımlarında dolandırıcılık tespit edilir. Fiziksel mağazalarda yapılan dolandırıcılıklarda aktif olarak kullanılabilir. Burada ki temel yapı banka müşterileri genel olarak kredi limitleri, maaşları oranında her ay ortalama bir alışveriş gerçekleştirirler. Bununla beraber müşterinin aldığı ürün yelpazesi de sürekli olarak benzerlik gösterebilir. Bütün bu değişken ve yapılar veri madenciliği kavramıyla incelenirse anlamsız olan bu verilerden anlamlı bilgi edilir ve bu bilgi sonucu dolandırıcı enlemini gerçekleştirme üzerindeyken alışverişini riskli olarak gruplayarak alışveriş işleminin gerçekleştirilmesi engellenebilir. Bunun için veri madenciliği kavramının farklı algortmaları üzerinden işlem yapılarak en iyi sonuçlar alınıp gerçek bir tahmin edici model oluşturulabilir.

7. VERİ MADENCİLİĞİ

Bu bölümde temel olarak veri madenciliği deyiminin, tarihçesi, süreçleri ve modelleri gibi konularda bilgi verilecektir.

7.1. Veri Madenciliği Tanımı

Veri madenciliği, en temel anlamıyla yüksek miktarda bilgiyi içeren sistemlerde yapılan bir ayrıştırma işlemidir. Ayrıştırmadan kasıt insanlar için gereksiz olduğu düşünülen verilerin birleştirilmesi işlemi ile yeni anlamlı veri (bilgi)nin oluşturulmasıdır. Oluşturulan bilgi çeşitli bilgisayar programları ile elde edilebilir. Veri madenciliği kelimesi yanlış algılanabileceğinden ötürü literatüre eşdeğer başka deyimlerde eklenmiştir. Bu deyimler;

- i. Knowledge mining from databases (Veri tabanlarında bilgi madenciliği)
- ii. Data / Pattern Analysis (Veri ve örüntü analizi)
- iii. Knowledge discovery from databases (Veri tabanlarında bilgi keşfi)

Veri madenciliği deyimine neden ihtiyaç duyulur kısmına gelinirse; günümüzde internet ve teknolojinin sürekli gelişimi ile dosyalama ve arşivleme işlemleri de bilgisayar ortamına aktarıldı. Bilgisayar sistemlerine geçen arşivleme insanlara bilginin daha güvenli saklanması, hızlı ve kolay erişim ve ekonomik açıdan birçok yarar sağladı. İnsanlar da tutulan kişisel veriler ve şirketler kayıtları işleyerek keşfedilmemiş birçok bilgiyi üretme şansına erişti. Bütün bu yararları düşünülerek günümüzde iş dünyasının farklı sektörlerinde veri madenciliği kavramı etkin olarak kullanılır (Bigus, 1996). Örnek vermek gerekirse;

- i. Pazarlamada uygulama alanları;
 - a) Alışveriş sepeti analizi

- b) Müşterileri değerlendirme işlemi
- c) Gelecek satışların tahmin edilmesi
- d) Müşteri alım alışkanlıklarını belirleme
- ii. Bankacılıkta uygulama alanları;
 - a) Kart dolandırıcılıkların tespit edilmesi
 - b) Harcama alışkanlıklarına göre müşteri gruplarının oluşturulması
 - c) Kredi notu belirleme işlemi
 - d) Kara para akışının tespiti
- iii. Sigortacılıkta uygulama alanları
 - a) Yeni poliçe isteyebilecek müşterilerin bulunması
 - b) Dolandırıcılıkların tespiti
 - c) Müşteri risk grubunun belirleme işlemi
- iv. Tıpta uygulama alanları;
 - a) Hastalıkların tanılarının belirlenmesi
 - b) Hastaya özgü tedavi oluşturma işlevi
- v. Sporda uygulama alanları;
 - a) Rakip taktik ve oynama alışkanlıkların belirlenmesinde
 - b) Maç için oyuncu performanslarının belirlenmesinde
 - c) Gelecek düşünülerek oyuncuların sakatlanma şansının belirlenmesinde

Veri madenciliği deyimiyile buluşan sektörler düşünüldüğünde büyük verilerin işlenmesi ve işlenen veriden oluşan bilgi şirketler için oldukça önemlidir. Bu yüzden veri madenciliği kalıbı pazarda güçlü kalma ve rakiplerle rekabet için kullanılması gereken güçlü bir stratejidir. Yapılacak bütün işlemler düşünüldüğünde veri madenciliği kavramının daha ayrıntılı açıklayabilmek için bir takım sorulara cevap vermeliyiz;

Veri Nedir?: Veri ham şekilde bulunan işlenmemiş küçük parçacıklara verilen isimdir. Veriler çeşitli ölçüm, deney ve araştırma yöntemleriyle elde edilir. Başka bir tanımla ifade etmeye çalışırsak bilginin oluşturulması için işlenmiş parçacıklardır.

Veri Nasıl Saklanır?: Verilerin saklanması geçmiş yıllar göz önünde bulundurulduğunda büyük farklılık göstermiştir. Eskiden veriler geleneksel olarak kâğıtlarda tutulmaktaydı. Teknolojinin ve internetin sürekli gelişimi sonucu veriler günümüzde veri tabanı yönetim sistemlerinde (VTYS) saklanır.

VTYS'nin avantajları:

- i. Veri güvenliğinin artırılması
- ii. Veri bütünlüğünü güçlendirme
- iii. Veri tekrarını azaltma

Veri Nasıl İşlenir?: Verinin işlenip bilgiye dönüştürülmesi için bazı işlem basamakları uygulanmalı ve veri işlenmeye hazır hale getirilmelidir.

7.2. Veri Madenciliğinin Tarihçesi

Veri madenciliği teknolojinin gelişimiyle doğrudan etkilenmiş ve şuan ki halini almıştır. İnsan yapısı düşünüldüğünde, her zaman verileri biriktirme, yorumlama ve bilgi edinme işlemlerini sırasıyla uygulamışlardır. Bu da bilgilerin geçmişten günümüze kadar sürekli taşınmasını sağladı.

Bilişim sektöründe uygulanmaya başlanan veri madenciliği deyimini ile bilgi keşfini düşünürsek tarihsel gelişim 1960 yılından itibaren başlar. 1960'ların sonlarına doğru veri tabanı sadece bilgilerin tutunduğu bir yapı olmaktan çıkmış ve üstünde işlem yapılabilmiştir. Yapılan işlem sonucunda da bilgiye ulaşılabileceği düşüncesi hakim oldu. 1970'lerde E.F.Codd tarafından yayınlanan ilişkisel veri tabanı makalesi ile insanlar veri madenciliği tanımı ile tanıştı. Aynı yıllardan ms-sql server ve Sybase gibi yapılar geliştirildi. 1980'lerde SQL standart dil olarak kabul edilerek birçok sektörde veri tabanı yönetim sistemi kullanımı arttı.

Veri madenciliği kavramının ortaya çıkışı 1990'lı yılları bulur. Oluşan bu yeni deyimle geleneksel yöntemler yerini, veri analizinin bilgisayar üzerinde çalışan algoritmalara bıraktı. Sonrasında bilim adamları veri madenciliği deyimine çeşitli yaklaşımlar getirmeye çalıştı. Aşağıdaki Şekil 7.1. 'de veri madenciliğinin gelişim süreci gösterilmektedir.

Evrimsel Aşama	İşletme Sorusu	Etkinleştirme Teknolojileri	Ürün Sağlayıcıları	Karakter Özellikleri
Veri Toplama (1960'lar)	'Son beş yıldaki toplam gelirim nedir?'	Bilgisayarlar, kasetler, diskler	IBM,CDC	Geriye döntük, statik veri teslimatı
Veri Erişimi (1980'ler)	'Geçen Mart New England'daki parça satışları nedir?'	Bağlantılı veri tabanları (RDBMS), Yapılandırılmış Sorgu Dili (SQL),ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Geriye döntük, kayıt düzeyinde dinamik veri teslimatı
Bilgi Depolama & Karar Desteği (1990'lar)	'Geçen Mart New England'daki parça satışları nedir?Boston'a derinleme.'	Çevrimiçi analitik işlem (OLAP), çok boyutlu veri tabanları, bilgi depolama	Pilot, Comshare, Arbor, Cognos, Microstrategy	Geriye döntük, çoklu seviyelerde dinamik veri teslimatı
Veri Madenciliği (Günümüzde Yükselen)	Gelecek aydaki Boston parça satışlarına ne olması muhtemeldir? Neden?	Gelişmiş algoritmalar, çok işlemcili bilgisayarlar, çok büyük veri tabanları	Pilot, Lockheed, IBM,SGI, Sayısız Girişim Şirketleri (yeni doğan sanayi)	Geriye döntük, etkin bilgi teslimatı

Şekil 7.1. Veri madenciliğinin gelişim süreci

Gelişen veri madenciliği şirketlerin ihtiyaçlarına göre farklılık gösterir. Bunlar;

Metin madenciliği: Haber, makale, web siteleri vb. yazısal kategorideki işlemleri gerçekleyen veri madenciliğine verilen isimdir.

Grafik madenciliği: Sık olarak model madenciliği olarak da bilinir. Grafikler üzerinde işlem yapılmasını sağlar. Günümüzde en çok sosyal ağ madenciliği alanında kullanılır.

Görsel madenciliği: İki veya üç boyutlu işlerde kullanılan veri madenciliğine verilen isimdir.

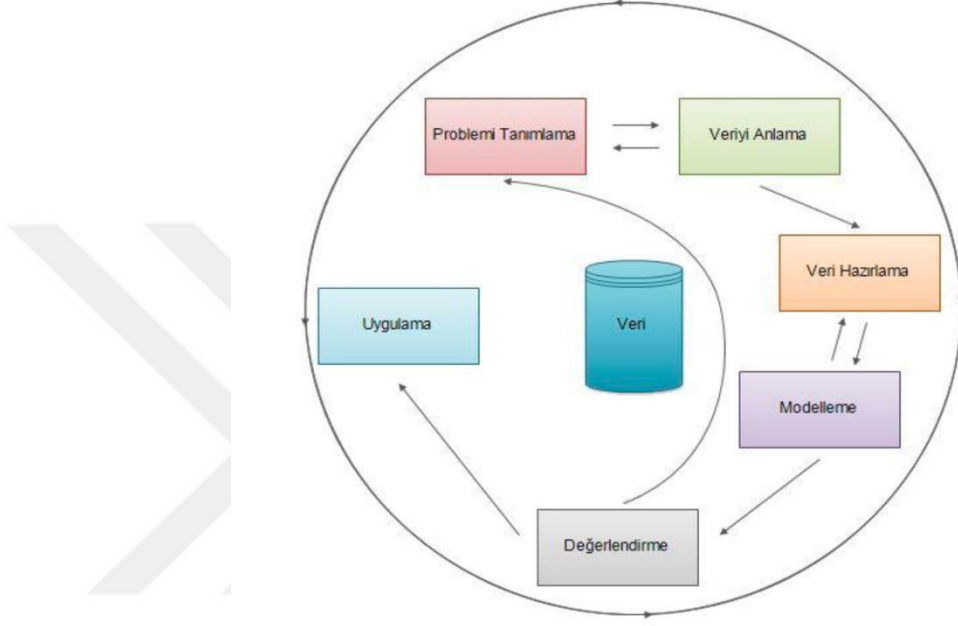
7.3. Veri Madenciliği Süreçleri

Başarılı bir veri madenciliği yapılması ve bilgi keşfi için bazı süreçler sırasıyla uygulanmalıdır. Bunlar 1996 yılında CRISP-DM (The Cross Industry Standart Process for Data Mining) konsorsiyumu tarafından belirlendi. Veri madenciliği süreci temel olarak altı adımda ifade edildi. Adımlar aşağıda ifade edildiği gibidir;

- i. Problemin tanımlanması işlemi
- ii. Verinin anlaşım işlemi

- iii. Verinin hazırlanması
- iv. Modelleme işlem basamakları
- v. Değerlendirme
- vi. Uygulama

olarak ifade edilir. Aşağıdaki Şekil 7.2. 'de veri madenciliği süreçleri gösterilmektedir.



Şekil 7.2. Veri madenciliği süreçleri

Problemin tanımlanması: Veri madenciliği süreçlerinin en önemlisidir. Doğrudan başarımı ve kaliteyi etkiler (amacı, uygulama alanı isterler, hedefler) doğru ve kesin olarak ifade edilmelidir. Problem tanımlanmasında yapılan en küçük bir hata bütün projeyi doğrudan etkileyecektir.

Verinin anlaşılması: Verinin anlaşılması oldukça önemlidir. Anlaşılmadan kasıt veriler üzerinde yer alabilecek sorunların belirlenmesi ve salt okunur hale getirilmesidir.

Verinin hazırlanması: Verinin hazırlanması oldukça zorlu ve veri madenciliğinin en fazla zaman kaybına sebep olan işlemidir. Bu basamakta temel olarak modelleme işlemine hazırlık (veri setinin oluşturulması), varsa problemlerin belirlenmesi ve düzenlenmesi işlemleri gerçekleşir. Yapılabilecek küçük bir hata modellemenin yanlış kurulmasına ve sonuç olarak yüksek miktarda vakit kaybına sebebiyet verir. Bütün

nedenler göz önüne alınarak bu süreç alt süreçlere bölünüp hata yüzdesi düşürülmeye çalışılmıştır.

Bu süreçteki alt süreçler aşağıda verilmektedir;

- i. **Toplama:** Giderilecek sorun için verilerin toplanacağı adımdır.
- ii. **Birleştirme:** Farklı kaynaklardan alınan verilerin ortak bir veri tabanında tutulma işlemidir.
- iii. **Karar verme:** Modele uygun veriye karar verilmesi işlemidir.
- iv. **Modelleme işlem basamakları:** Tanımlanması bitirilmiş çözülecek problem için en uygun modelin belirlenmesi işlemidir. Model belirlenerek üzerinde madencilik yöntemleri gerçekleştirilir. Modeller tanımlayıcı ve tahmin edici olarak ikiye ayrılır.
- v. **Değerlendirme:** Model değerlendirilerek eksikleri varsa giderilir. Uygulama öncesi hatasız hale getirilmeye çalışılır. Bu aşamada önceki süreçleri kontrol işlemleri de tekrar sağlanabilir.
- vi. **Uygulama:** Oluşturulan ve hatasız olduğu düşünülen modelin uygulandığı son veri madenciliği sürecidir.

7.4. Veri Madenciliği Modelleri

Veri madenciliği modelleri temel olarak iki başlık halinde incelenebilir. Bunlar;

- i. Tanımlayıcı modeller (birliktelik kuralları, kümeleme)
- ii. Tahmin edici modeller (sınıflandırma, regresyon, zaman serisi analizi)

7.4.1. Tanımlayıcı Modeller

Tanımlayıcı modelde karar vermeyi sağlayacak örüntülerin tanımlanması amaçlanmıştır. Model, kümeleme analizi ve birliktelik kuralları başlıkları altında incelenir.

7.4.1.1. Kümeleme

Kümeleme analizi temel olarak birbirlerine benzeyen nesnelere birleştirilerek ayırma işlemidir. Oluşan kümelerde küme içi benzerlik maksimum, kümeler arası benzerlik minimum olmalıdır. Benzerliğin veriler arasındaki uzaklıklar ile ifade

edebiliriz. Uzaklığı Öklid, Manhattan, Minkowski bağıntıları ile hesaplayabiliriz. Kümeleme algoritmaları genellikle çağrı merkezleri gibi büyük verilerin tutulduğu uygulamalarda kullanılır. Kümeleme analizi iki başlık halinde toplanabilir.

Bunlar;

- i. Hiyerarşik Yöntemler
- ii. Hiyerarşik Olmayan Yöntemler

Hiyerarşik Yöntemler: Hiyerarşik yöntemlerde veri tabanında bir küme ağacı oluşturulup ağaç yapısı için öncelikle ana küme belirlenir belirlenen küme alt kümelere ayrılarak aşamalı ağaç benzeri bir yapı meydana getirir. Bu yöntemde en fazla kullanılan algoritmalar;

- i. En yakın komşu algoritması
- ii. En uzak komşu algoritması

Hiyerarşik Olmayan Yöntemler: Hiyerarşik olmayan yöntemlerde kümeler, önceden belirlenen sayıda ayrılarak gerçekleşir. Yöntemde en fazla kullanılan algoritmalar;

- i. K- Ortamalar Yöntemi
- ii. Metoid Parçalama Yöntemi
- iii. Yığma Kümeleme Yöntemi

7.4.1.2. Birliktelik Kuralları

Olayların beraber gerçekleşme ihtimallerinin çözümler. Ürünler arasındaki ilişkiler belirlenir ve bunların eşzamanlı yapılma olasılıkları hesaplanır. En yoğun kullanılan birliktelik kuralı algoritması apriori 'dir. Yaygın olarak alışveriş sepetlerinin analizinde kullanılır. Analizler sonucu müşteri alışkanlıkları belirlenerek raf dizaynı, market içinde ürünlerin yerleştirilmesi gibi olaylarda aktif olarak kullanılır. Bir örnek ile ifade edersek markette şampuan ve parfüm alanların %70'i aynı anda cilt kremi almıştır. Buradan yola çıkarak market sahipleri müşterilerinin alışkanlığını tespit ederek reyon ve raf tasarımını geliştirip satış rakamlarını yukarı çekebilirler.

7.4.2. Tahmin edici Modeller

Tahmin edici modeller isminde de anlaşılabilir gibi veritabanı yönetim sistemi üzerinde ki parçacıklardan tahminler yapmaktadır. Model üç alt başlıkta gerçekleştirilir. Bunlar iki grupta ele alınabilir;

- i. Regresyon
- ii. Sınıflandırma

7.4.2.1. Regresyon

Bir bağımlı veya birden çok bağımsız değişken arasında ki ilişkinin matematik yardımıyla ifade edilmesine verilen isimdir. Tahmin edici model olup veriler üzerinde tahmin işlemini benimser. Bağımlı ve bağımsız değişken arasındaki yakınlık korelasyon kat sayısı ile bulunmalıdır. İlişkinin yapısına göre iki farklı şekilde açıklanır.

- i. Doğrusal Regresyon Modeli
- ii. Doğrusal Olmayan Regresyon Modeli

Daha kolay ifade etmek gerekirse iki veya daha fazla değişken arasında neden sonuç ilişkisi bulunan yapı üzerinde gerekli görülen tahmin ve kestirmelerin yapılması işlevine verilen isimdir. Örnek olarak kısa neden – sonuç ilişkileri verilirse (Gelir – harcama, Çalışma zamanı – Alınan not, vb.)

8. SINIFLANDIRMA YÖNTEMİ

Sınıflandırma tahmin edici bir model olup veri madenciliğinin en sık kullanılan yöntemlerinin başında gelir. Sınıflandırma yapısının asıl amacı veri tabanı yönetim sistemindeki gizli örüntülerinin bulunup ortaya çıkarılmasıdır. Veriyi sınıflandırmak için bazı alt işlemler gerçekleştirilmelidir. Bu alt işlemler;

- i. Veri tabanının bir kısmı eğitim amacıyla kullanılır. Eğitim sonucunda sınıflandırma kuralları oluşturulur.
- ii. Oluşturulan kurallar yardımıyla gerçekleşebilecek yeni durumlara karşı hangi kararların verileceği belirlenir.

Sınıflandırma yöntemini bir örnek ile ifade edersek; (Bankacılık sektöründe bir müşteri kredi talebiyle bankaya giderse, bankalar müşterinin gerekli değerlerini göz önünde bulundurarak müşterisinin risk durumunu iyi veya kötü olarak sınıflandırabilir.) Aynı zamanda sınıflandırma yöntemi tıp sektöründe de oldukça fazla kullanılır. Örneğin; kanser hastalığına sebep olan hücrelerin seçiminde. Sınıflandırma birçok teknik ile gerçekleştirilebilir (Berry, Linoff, 2004). Bu teknikler;

- i. En yakın k- komşu algoritması
- ii. Yapay sinir ağları ile sınıflandırma
- iii. Karar Ağaçları ile sınıflandırma

8.1. En Yakın k - Komşu Algoritması ile Sınıflandırma

Sınıflandırma yöntemleri arasında kullanılan tekniklerden biride en yakın k- komşu algoritmasıdır. Tekniğin uygulanması kolay ve gözetimli bir öğrenme algoritmasıdır. T.M. Cover ve P.E Hart tarafından 1967 yılında ortaya atılmıştır. Sınıfları daha önceden belirtilmiş bir veri setinde sonradan eklenecek verinin sınıfının

belirlenmesi gibi işlemlerde kullanılır. Bu yaklaşımda yeni eklenen verinin bütün değerlere olan uzaklıkları belirlenir. Uzaklıklar üç tip fonksiyonla hesaplanır. Bunlar;

- i. Öklid
- ii. Manhattan
- iii. Minkowski

Sınıflandırma yönteminde en yakın k-komşu algoritması ile yeni değer in sınıfının belirlenmesi için kullanılan adımlar;

- i. k parametresi belirlenir. Belirlenen k parametresi genellikle tek sayı olarak oluşturulur. Tek olmasının amacı hiçbir zaman iki sınıfa eşit sayıda yakınlığın olma ihtimalinin ortadan kaldırılmasıdır.
- ii. Yeni değer in en yakın komşuları belirleneceği için tüm veri setine olan uzaklıklar hesaplanır.
- iii. Uzaklıkları en yakın k tanesi seçilir.
- iv. Seçilen k kadar değişkenini sınıflarına bakılarak en fazla tekrarlayan sınıf yeni değer in de sınıfı olarak atanır.

8.2. Yapay Sinir Ağları ile Sınıflandırma

Yapay sinir ağları en basit anlatımıyla insan beyni çalışma yapısının modellenerek bilgisayar ortamına aktarılmasıdır. Aktarımın gerçekleştirilmesi için insan sinir hücresinin çalışmasını incelenerek bilgisayarda yapay sinir hücresi oluşturulur. 1890 yılında yapay sinir ağları tekniğiyle ilgili çalışmalar yapılmaya başlandı. Yapay sinir ağlarının temel uygulama alanları aşağıda belirtildiği gibidir;

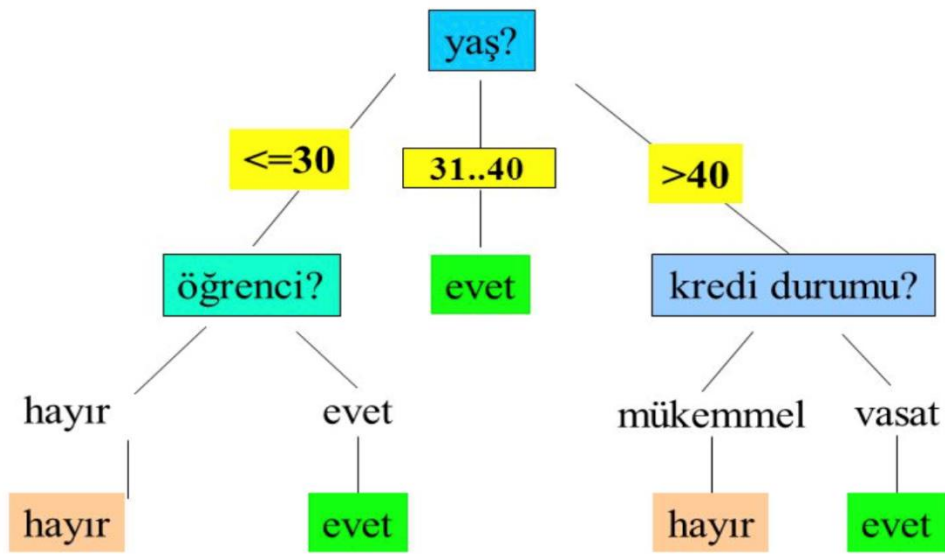
- i. Robotik
- ii. Tahmin yürütme
- iii. Optimizasyon
- iv. Sistem modelleme
- v. Sınıflandırma

YSA, karar ağaçları ile beraber en çok kullanılan sınıflandırma tekniklerinden biridir. Yapay sinir ağları ile sınıflandırmada veriler ortak özellikleri baz alınarak sınıflandırılır (Temel, Çamdeviren, Akkuş, 2005).

8.3. Karar Ağaçları ile Sınıflandırma

Sınıflandırma modelini sağlamak için birçok yöntem seçilmiştir. Fakat sınıflandırma da en çok kullanılan yöntem karar ağaçları ile sınıflandırmadır.

Karar ağaçları ile sınıflandırma yapısını biraz açılırsa bu yapı akış şemaları şeklindedir. Bütün öznitelikler birer düğümü simgeler. Geri kalan dal ve yapraklar karar ağacının elemanlarıdır. Yaprak ve kök arasındaki bütün yapılar dal olarak adlandırılır. Aşağıdaki Şekil 8.1. 'de karar ağacının yapısını anlatan görsel bulunmaktadır.



Şekil 8.1. Örnek karar ağacı yapısı

Karar ağaçlarının uygulanması kısmında bir takım algoritmalarından yararlanılabilir. Bu algoritmalar aşağıdakilerdir;

- i. ID3 Algoritması
- ii. C4.5 Algoritması
- iii. MARS (Sayısal verilerin iyi işlenebilmesi için kullanılır.)
- iv. Random forest (Birden fazla karar ağacı kullanılarak sınıflandırma kalitesi yükseltir.)
- v. Boosted Trees (İçinde hem sınıflandırma hem de regresyon problemlerinin çözümünü barındırır.)
- vi. Chi-Kare (Birden fazla bölme işlemine izin verir.)

Karar ağacı algoritmaları

i. **ID3 algoritması:** 1983 yılında J.R. Quinlan tarafından geliştirilen bir algoritmadır. ID3 algoritması önceden belirlenmiş veri seti üzerinden karar ağacı tekniği kullanılarak elde edilmiştir. Algoritmanın temeli entropi ve bilgi kazanımı terimleri üzerinedir.

Entropi: Sistem üzerinde ki belirsizliğin ifade edilmiş şeklidir.

Bilgi kazanımı: Entropinin tersi olmakla birlikte sıfır ve bir arasındadır. Sonradan verilen değer için sınıflandırma sonucunu kazanacağı değer miktarıdır.

ii. **C4.5 Algoritması:** 1993 yılında J.R. Quinlan tarafından ID3 algoritması üzerinde ki eksiklerin giderilmesi için geliştirilen bir algoritmadır. Temeli ID3 üzerine kuruludur. C4.5 algoritmasının en önemli yeteneklerinden biri sayısal değerler üzerinde de çalışabilmesidir. Karar ağacı tekniğinin bu özellikleri baz alınarak kullanımı en fazla olan algoritmasıdır. Tıpkı ID3 algoritması gibi temelleri entropi ve bilgi kazanımı üzerinedir (Silahtaroglu, 2008).

Karar Ağaçları ile Sınıflandırma Avantajları

Karar ağacı ile sınıflandırma tekniğinin sağladığı avantajlar aşağıdadır. Bu avantajlar;

- i. Karar ağacını yorumlamak kolay ve zahmetsizdir.
- ii. Anlaşılabilir kurallar oluşturulabilir.
- iii. Ayrık nitelik değerleri için kullanılır.
- iv. Çoklu çıktı probleminin önüne geçer.
- v. Doğrusal olmayan parametre yapısı ağacın performansını büyük ölçüde etkilemez.

Karar Ağaçları ile Sınıflandırma Dezavantajları

Karar ağacı ile sınıflandırma tekniğinin sağladığı dezavantajlar aşağıdadır. Dezavantajlar;

- i. Yapı küçük ağaçlar için kolay, karmaşık ağaçların okunması için karmaşıktır.
- ii. Karar ağacına uygulamadan önce veri seti düzenlemelidir (Özkan, 2016).

9. GEREÇ ve YÖNTEM

Bu bölümde temel olarak yapılmış olan projede tercih edilen teknolojiler ve diller hakkında bilgiler verilmekte olup ayrıca hangi nedenlerden dolayı tercih edildikleri anlatılmaktadır.

9.1. C4.5 Algoritması

C4.5, veri madenciliğinde kullanılan sınıflandırma modeli içindeki karar ağacı ile sınıflandırma tekniğinin bir algoritmasıdır. Karar ağacı ile sınıflandırma tekniğinin ID3 algoritması ile beraber kullanımı en fazla olan algoritmasıdır. J.R. Quinlan tarafından 1993 yılında ID3 algoritması üzerindeki eksikleri gidermek için geliştirilerek veri madenciliği ortamına kazandırılmış bir algoritmadır. C4.5 algoritmasının insanlara kazandırdığı en önemli ayrıcalık sayısal veriler üzerinde işlem yapabilme yetisi ve bu yeti sayesinde karar ağacımızı düzenleyebilme imkanındır. Bununla birlikte kayıp değerler ile baş edebilme ve bölünme bilgisi gibi ek özellikleri de kullanıcılarına sunar. Tabi ki bu algorithmada üzerine geliştirilen ID3 algoritması gibi entropi ve bilgi kazanımı kavramları üstüne kurulmuştur.

Entropi: Düzensizliğin sistem üzerinden ifade edilmesine verilen isimdir. Karar ağacı tekniğinde ki algoritmalarda kullanımının asıl nedeni dallanma işlemine nerden başlanabileceğinin tespitidir. Kısacası kök hücrenin ortaya çıkarılmasında kullanılır. Karar ağacı, kök hücreden dallanmasını devam ettireceği için entropi hesabıyla kök belirlemek oldukça elzemdir. Ayrıca bu terim bir termodinamikçi olan Shannon tarafından geliştirilmiştir. Entropi hesabını yaparak kök değeri ve dallanmanın belirlenmesi için kullanılacak formül aşağıdaki (9.1.) 'deki gibidir.

$$\sum_{i=1}^n P(x_i) \log_2 \frac{1}{P(x_i)} = - \sum_{i=1}^n P_i \log_2 p_i \quad (9.1.)$$

Formül üzerinde ki n değeri entropisi hesaplanacak nitelik değeri içinde ki ihtimal sayısını, p_i değeri ise ihtimalin toplama olan oranının ifade etmektedir. Tüm nitelikler üzerinden yapılan entropi hesabına göre dallanma, entropisi en fazla olan değerden başlar.

Bilgi kazanımı: ID3, C4.5 gibi algoritmalarda ayırt edici niteliğin tespit edilmesi için gerçekleşen eylemdir. Bu eylem entropi hesabına dayalıdır. Fakat entropi ile bağlantısı terstir. Bilgi kazanım değerleri sıfır ile bir arasında olmakla beraber bilgi kazanımı bire yaklaştıkça niteliğin önemi azalırken sıfıra yaklaşma durumunda önemi artar (<http://bilgisayarkavramlari.sadievrenseker.com.>, Erişim Tarihi: 21 Nisan 2019)

$$\text{Gain}(T, X) = \text{Entropy}(T) - \text{Entropy}(T, X) \quad (9.2.)$$

Bilgi kazanımında sınıf değeri ile beraber tüm nitelikleri entropi değerleri hesaplanır. Bilgi kazanımını hesaplamak için sınıf değerinin entropisinden, sınıf değeri ile her niteliğin ağırlıklı entropisi hesaplanarak çıkarılır. Formülde ki T, sınıf değeri X ise bir niteliktir. Bilgi kazanımı en yüksek olan değerden ağaç dallandırılır.

C4.5 algoritmasının adımları;

- i. Belirlenen veri seti üzerinde ki sınıf ve nitelik değerleri için entropi değerleri hesaplanır. Entropi değerleri yukarıda ki formül ile gerçekleştirilir.
- ii. Entropisi hesaplanan sınıf değeri ile diğer nitelikler arasında bilgi kazanımları hesaplanır.
- iii. Bilgi kazanımı en iyi olan değer kök olarak seçilerek dallanma kök üzerinden başlatılır.
- iv. Niteliklerin yerleştirimi tamamlanıncaya kadar ilk üç işlem tekrar edilir.

Karar ağacı ile sınıflandırma tekniğinin bizlere kazandırdığı (anlaşılabilir kurallar, okunabilirlik kolaylığı, vb.) ve bu tekniğin algoritması olan C4.5 algoritması ile hem sayısal hem de kategorik veriler üzerinde çalışabilme imkânı gibi avantajlar projede kullanımını sağladı.

9.2. Weka Yazılımı

Yeni Zelanda Waikato Üniversitesi tarafından makine öğrenimini gerçekleştirmek amacıyla geliştirilmiş bir yazılımdır. Weka kelimesi “Waikato Environment for Knowledge Analysis” kelimelerinin baş harfleri baz alınarak oluşturuldu. Günümüzde yazılan makale ve tezlerde (makine öğrenmesi, veri madenciliği) en fazla kullanılan yazılımlardan biri halini almıştır. Popülerliğini Java dilinde oluşturulması ve kütüphanenin .jar dosyaları halinde olmasına borçludur. Çünkü Java dili günümüz dünyasında en aktif kullanılan yazılım geliştirme dilidir. Weka ‘nın popülerliğinin bir başka sebebi ise tüm işletim sistemlerine (Windows, macOS, Linux) uygun sürümlerinin bulunmasıdır.

Dosya uzantısı olarak .arff (Attribute Relationship File Format) kullanır. Buradan anlaşabileceği gibi veri setini .arff uzantılı şekle getirerek Weka ortamına aktarabiliriz. Veri madenciliği işlemlerinden kullanılan .csv uzantısının da .arff formatına çevirme işlemi yazılımın kendi içeriğinde mevcuttur (Şeker, 2013).

Weka tüm iş sektörlerinde aktif olarak kullanılabilir. Anlamsız veriden anlamlı veriyi oluşturabilir. İş zekâsı, veri analizi gibi uygulamalarda kendisinden faydalanımı yüksektir. Weka ile aşağıdaki işlemler gerçekleştirilebilir;

- i. Sınıflandırma (Classification)
- ii. Kümeleme (Clustering)
- iii. Birliktelik kuralları (Association Rules)

9.2.1. Weka 'da Kullanılan Ekranlar

Weka uygulamasını kendi sistemimize uygun olarak indirip yükleme işleminin gerçekleştirildiği takdirde kullanıcıları Weka logosu, ismi ve applications başlığı altındaki 5 seçenek karşılar. Açılış sayfası ekranını Şekil 9.1. ‘dedir. Bu seçenekler;

Explorer: Weka üzerinde daha önceden eklenmiş data klasörü altında ki veri setleri veya kendi oluşturduğumuz veri seti üzerinde işlem yapılabilen kısımdır. Hazır veya sonradan oluşturulan veri seti üzerinde (sınıflandırma, kümeleme, ilişkilendirme) gibi işlemlerinin yapılması sağlar.

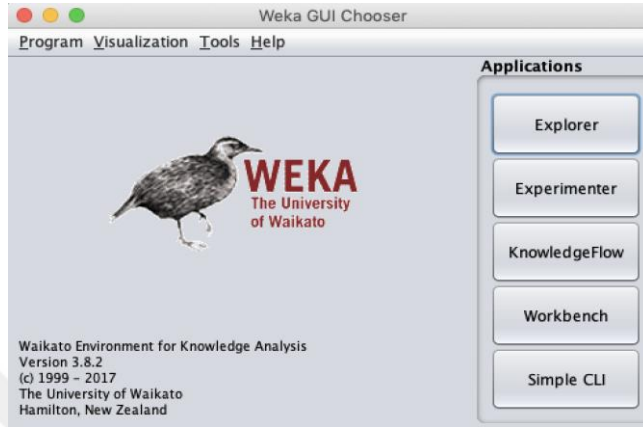
Experimenter: Kullanıcının şemalarını işlerken olması gerekenden daha iyi deneyler oluşturularak analiz etmesini sağlar.

Knowledge Flow: Amacı kullanıcının Weka üzerinde çok sayıda bulunan kütüphane görsel olarak erişim sağlanmasıdır.

Workbench: Üzerinde çalışılan ortam sekmesidir.

Simple CLI: Terminal tarzı bir sistemdir

(<http://bilgisayarkavramlari.sadievrenseker.com.>, Erişim Tarihi: 17 Nisan 2019).



Şekil 9.1. Weka açılış ekranı

Yapımı tamamlanan tez de Explorer sekmesi altında ki işlemler gerçekleştirilerek tez gerçekleştirilmeye çalışıldı. Explorer sekmesine tıkladığı vakit üst menü 6 başlık halindedir (<https://www.eecs.yorku.ca.>, Erişim Tarihi: 25.04.2019). Bu başlıklar aşağıdaki gibidir;

Preprocess: Veri setini yüklemeyen ve yapılacak işlem gerçekleştirilmeden önce kullanılan bir ön hazırlık sürecidir. Bu adımda oluşturulan veri setini veya önceden Weka 'nın içinde bulunan data klasöründen seçilen .arff uzantılı dosya sisteme yüklenebilir. Yükdikten sonra arayüz üzerinden veri setinin nitelikleri ve nitelik içerisinde ki değerleri görülebilir. Arayüz üzerinde nitelik silme işlemi de yapılabilir. Ön hazırlık sayfası aşağıdaki Şekil 9.2. 'de gösterilmiştir.

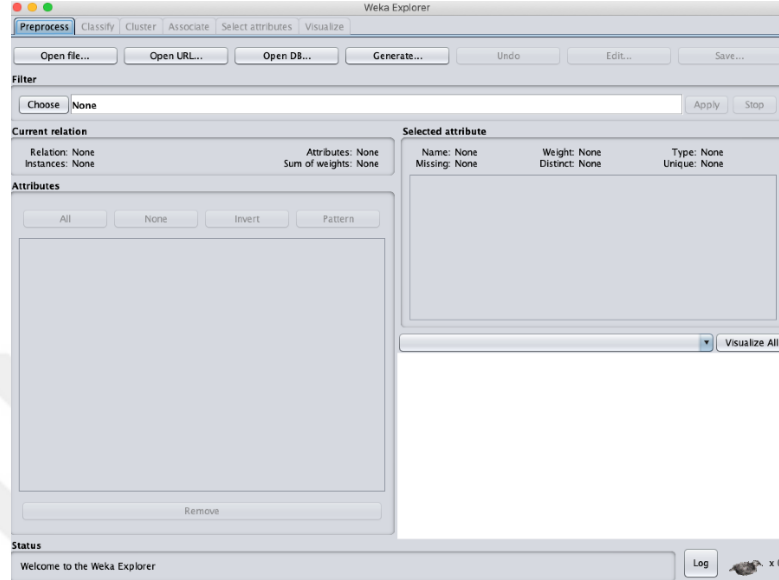
Classify: Yüklü ön hazırlığı yapılmış veri seti üzerinde kullanılmak istenilen algoritmalarla sınıflandırma işleminin gerçekleştirildiği yerdir.

Cluster: Ön hazırlık işlemleri gerçekleştirilmiş veri seti üzerinde kullanılmak istenen algoritmalarla kümeleme işleminin yapılabildiği sayfadır.

Associate: Ön hazırlığı yapılmış veri seti üzerinde verinin istenen algoritmayla yorumlanıp birliktelik kurallarının oluşturulduğu yapıdır.

Select Attributes: Yüklendiği veri seti üzerinde ki birbirleri ile ilişkili verilerin niteliklerinin belirlendiği sayfadır.

Visualize: Niteliklerin birbirleri ile olan ilişkisinin grafik olarak ifade edilebildiği yapıdır.



Şekil 9.2. Weka ön hazırlık sayfası

Tez kapsamında sınıflandırma yönteminin içindeki tekniklerden karar ağacı ile sınıflandırma tekniği ve C4.5 algoritması kullanılmıştır. Sınıflandırma yöntemi ile proje geliştirildiği için Classify sekmesi üzerinde ayrıntılı durulacaktır.

Classify (Sınıflandırma)

Ön hazırlık işlemi tamamlandıktan sonra ve veri seti sınıflandırma yöntemine uygun hazırlandıktan sonra classify sekmesi üzerinden choose butonuyla uygulanmak istenen algoritma seçilir. Seçim sonrası test options üzerindeki seçeneklerden biri seçilerek sınıflandırma işlemi yapılabilir. Yapılan seçim sonucu oluşan sınıflandırma classifier output kısmında kullanıcıya aktarılır. Test options bölümü üzerinden 4 farklı işlem gerçekleştirilebilir. Bunlar;

- i. Use training set: Veri setinin ne kadar doğru sınıflandırıldığı kontrol eden butondur.
- ii. Supplied test set: Daha önceden sınıflandırılan veri seti üzerine yeni yükleyecek verinin ne kadar doğru sınıflandırıldığını kontrol eder.

- iii. Cross – validation: Folds kısmına girilen yüzdelik kısmını test ederek geriye kalan yüzde kadar eğitim işlemi yapar.
- iv. Percentage split: Verilecek yüzdeye göre sınıflandırmanın iyiliğini tespit eder.

Gereken tüm işlemler yapıp (algoritma ve test options seçimi) sınıflandırma işlemi yapıldığı takdirde classifier output altında kullanıcının göreceği bazı terimlerin anlamları aşağıda belirtilmiştir;

Number of Leaves: Sınıflandırılan ağaçta ki yaprak sayısıdır.

Size of Tree: Sınıflandırılan ağacının boyutunun gösterildiği kısımdır.

Time taken to build model: Modelin oluşturulması için harcanan zamanın kullanıcıya aktarıldığı kısımdır.

Time taken to Test Model on Training Data: Eğitim verisi üzerindeki test modeli için harcanan zamanın gösterildiği kısımdır.

Correctly Classified Instances: Sınıflandırma sonucu oluşturulmuş modeldeki eğitim veri setinden kaçının doğru sınıflandırıldığını gösterildiği kısımdır.

Kappa Statistic: Kappa testi sonucu oluşturulan değer gösterildiği yerdir. Değer sıfır ve bir arasında olmak zorundadır. bire yakınlık uyumu temsil eder.

Mean Absolute Error: Sınıflandırma sonucunda ki mutlak hata oranını gösteren kısımdır.

Root Mean Squared Error: Karekök ortalama hatasını gösterir.

Relative Absolute Error: Sınıflandırma sonucu göreceli mutlak hatayı gösterir.

Root Relative Squared Error: Kök bağı kare hatasının gösterildiği yerdir.

Total Number of Instances: Veri setine girilen toplam eğitim verisinin sayısının gösterildiği kısımdır.

Detailed Accuracy By Class: Sınıflandırma işlemi sonucu oluşan sınıflara göre oluşan detaylı doğruluk oranını gösterir.

Confusion Matrix: Sınıflandırma sonucu oluşan karmaşıklık matrisinin gösterildiği terimdir (<https://ceaksan.com/tr.>, Erişim Tarihi: 20 Nisan 2019).

Confusion Matrix tarafından üretilen sonuçların yorumlanması Weka için oldukça önemlidir. Önemi düşünülerek karmaşıklık matrisinin içerisindeki terimler kısaca açıklanacaktır (<https://veribilimcisi.com.>, Erişim Tarihi: 14 Mayıs 2019).

True Positive (TP): Doğru olmasını düşündüğümüz sonucun doğru olması durumudur. Düşündüğümüz sonucun algoritma ile de ispatıdır.

True Negative (TN): Yanlış olmasının düşünülen durumun algoritma ile de yanlış olmasıdır.

False Positive (FP): Yanlış olması düşünülen durumun algoritma ile doğru olarak gösterilmesine verilen isimdir.

False Negative (FN): Doğru olan durumun yanlış olarak gösterilmesidir.

Karmaşıklık matrisi doğruluk oranları sonucu ortaya bazı oranlar çıkmaktadır. Oranların tanımı ve formülleri aşağıda ki gibidir;

Accuracy Rate (Doğruluk Oranı): Oluşturulan algoritma baz alınarak oransal olarak doğruluğun ifade edildiği şeklidir. TP ve TN değerlerinin toplam orana bölünmesi ile elde edilir. Toplam değeri, karmaşıklık matrisindeki tüm değerlerin (TP, TN, FP, FN) toplamıdır. Doğruluk oranı formülü aşağıda (9.3.) 'de gösterilmektedir.

$$\text{Accuracy Rate} = (\text{TP} + \text{TN}) / \text{Toplam} \quad (9.3.)$$

Misclassification Rate (Yanlış sınıflandırma oranı): Oluşturulan algoritmanın hangi oranda yanlış değer gösterdiğini ifade eder. FP ve FN değerlerinin toplamının toplam orana bölünmesi ile elde edilir. Toplam değeri, karmaşıklık matrisinde ki tüm değerlerin (TP, TN, FP, FN) toplamıdır. Yanlış sınıflandırma oranı formülü aşağıda (9.4.) 'de gösterilmektedir.

$$\text{Misclassification Rate} = (\text{FP} + \text{FN}) / \text{Toplam} \quad (9.4.)$$

Sensitivity Rate (Duyarlılık oranı): Asıl pozitif değerlerin oranı olarak da bilinir. TP nin var olan doğrulara bölünmesi işlemidir. Var olan doğrular veri setindeki toplam doğrulardır. Duyarlılık oranı formülü aşağıda (9.5.) 'de gösterilmektedir.

$$\text{Sensitivity Rate} = \text{TP} / \text{Var olan doğrular} \quad (9.5.)$$

False Positive Rate: Yanlış olması gereken durumların ilgili algoritma sonucu doğru çıkan değerlere oranıdır. False positive değerinin gerçek yanlışlara bölümü ile ifade

edilir. Gerçek yanlış terimi ilgili veri seti üzerindeki toplam yanlışları anlatmaktadır. False positive rate formülü aşağıda (9.6.)’da gösterilmektedir.

$$\text{False Positive Rate} = \text{FP}/\text{Gerçek yanlışlar} \quad (9.6.)$$

True Negative Rate: Tahmini yanlış değerlerin asıl yanlışlara bölünmesine verilen isimdir. Asıl yanlışlar veri seti üzerinde ki tüm yanlışlardır. True negative rate formülü aşağıda (9.7.) ‘de gösterilmektedir.

$$\text{True Negative Rate} = \text{TN}/\text{Asıl yanlışlar} \quad (9.7.)$$

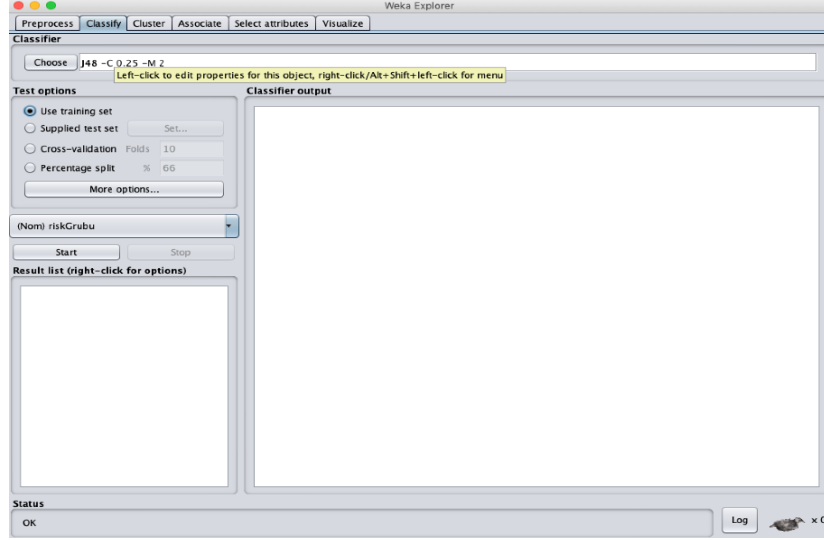
Precision (Hassasiyet): Sınıflandırma algoritması sonucu bütün sınıfların doğruluk oranlarının kullanıcıya aktarıldığı değerdir. Hassasiyet oranı formülü aşağıda (9.8.) ‘de gösterilmektedir.

$$\text{Precision} = \text{TP}/(\text{TP}+\text{FP}) \quad (9.8.)$$

Prevalance (Yaygınlık): Sınıflandırma sonucu doğruluğun sıklığının gösterildiği değerdir. Asıl doğrular değerinin toplam değerine bölümüyle elde edilir. Asıl doğrular ilgili veri setinde ki tüm doğruları, toplam değeri ise (TP+TN+FP+FN) işlemini ifade eder. Yaygınlık oranı formülü aşağıda (9.9.) ‘da gösterilmektedir.

$$\text{Prevalance} = \text{Asıl doğrular} / \text{toplam} \quad (9.9.)$$

Weka yazılımının arayüz kolaylığı veri madenciliği işlemleri üzerinde ki popülaritesi ve başarımı, çalışmada kullanılan sınıflandırma modelinin C4.5 algoritmasını içerisinde barındırmasını ve classifier output ekranının içeriği projede kullanımımı etkileyen en önemli faktörlerdir. Ayrıca C4.5 algoritması Weka yazılımı içerisinde J48 ismini almaktadır. Weka ‘nın sınıflandırma ve sonuç ekranı aşağıdaki Şekil 9.3. ‘de gösterilmiştir.



Şekil 9.3. Weka sınıflandırma ve sonuç ekranı

9.3. PHP

Dinamik biçimde olan web tabanlı sayfaların yapım aşamasında kullanılmakta olan bir programlama dili olarak adlandırılır. PHP ‘nin açılımı ele alınacak olunursa Hypertext Preprocessor olarak tanımlanmaktadır. Php, ilk olarak 1995 yılında yazılım geliştirme uzmanı Rasmus Lerdorf ‘un yapmış olduğu çalışmalar sonucu ortaya çıkmıştır. Yazılım uzmanının Php programlama dilini ortaya ilk olarak çıkarmasındaki amacı internet tabanlı sitelere giren ziyaretçi sayılarını belirlemektir. Ancak sonra Php dili üzerinde yapılan araştırma ve geliştirme sonucunda Php dili açık kaynak kodlu bir dil olarak piyasalara sunulmuştur ([https://www.mediaclick.com.tr.](https://www.mediaclick.com.tr), Erişim Tarihi: 11 Mart 2019).

Php ‘nin tezde kullanımını aşağıdaki avantajlar sağlamıştır:

- i. En temel avantajı açık kaynak kodlu olması ve herkes tarafından geliştirilebilmesi
- ii. Sanal sunucular yardımı ile tüm işletim sistemlerinde çalışabilme kabiliyeti
- iii. Php çalışma motor kodlarının hızlı yorumlaması sayesinde hızlı olması
- iv. Çok sayıda kaynağa sahip olması
- v. Derleme için özel bir teknolojiye gerek olmaması, Notepad vb. uygulamalarda kolayca derlenmesi
- vi. Sanal veya gerçek sunucuda kolay bir şekilde çalışabilmesi vb.

9.4. HTML

HTML, aslına bakıldığında vakit bir programlama dili değil, işaretleme dili olarak adlandırılmaktadır. HTML adı altında yazılan işaretleme kodları web browserlar aracılığı ile web tabanlı uygulamaların görsel olarak görünümü sağlanmaktadır. HTML, görsel görünümün yanında metin, video vb. görsel öğelerin web tabanlı uygulamalar üzerinde istenildiği gibi yerleştirme veya konumlandırma işlemi gerçekleştirebilir. Ayrıca arama motorlarına web tabanlı uygulamalar hakkında bilgi verilmesini sağlamaktadır. HTML dili yazmak için bilgisayara ayrı olarak bir uygulama veya derleyici yüklemeye gerek yoktur. Hali hazırda var olan notepad, wordpad vb. uygulamalar HTML dilini yazmak için yeterlidir. HTML 'in açılımı ele alınacak olunursa; Hyper Text Markup Language yani Zengin metin işaretleme dili olarak adlandırılmaktadır (<https://www.biltektasarim.com.>, Erişim Tarihi: 11 Mart 2019).

HTML 'in kullanımı aşağıdaki avantajları düşünülerek gerçekleştirildi;

- i. HTML işaret dilinin öğrenilmesinin kolay olması
- ii. Tüm web browserlarda çalışabilmesi
- iii. Daha az kaynak kullanarak yapılması istenilen her isteğin hızlı bir şekilde gerçekleştirilmesi
- iv. Bütün işletim sistemlerinde çalışabilir olması vb.

9.5. CSS

Tarihe bakıldığında vakit ilk olarak Tim Barners-Lee ile çalışan Hakon Wium un yapmış olduğu çalışmalar sonucunda 1994 yılında öneri olarak sunmuştur. Daha sonra 1996 yıllarında Dünya Çapındaki Ağ Birliği yani W3C tarafından gerçekleştirilen çeşitli çalışmalar sonucunda CSS1 yani ilk CSS piyasaya sunulmuştur. İnternet kullanımının artması ile birlikte yeni ihtiyaçlar ortaya çıkması W3C 'nin ilk CSS 'i geliştirerek 1998 yılında ikinci CSS olan CSS2 yi piyasa sunulmuştur. CSS3 için günümüzde hala çalışmalar devam etmektedir. CSS 'in açılımı Cascading Style Sheet Türkçe anlamı Basamaklı Stil Şablonları olarak adlandırılmaktadır.

CSS, basit olarak açıklanacak olunursa web tabanlı uygulamaların görselliğini ayarlamak veya oluşturmak için kullanılan her türlü yapıyı düzenlemek için

kullanılmakta olan bir biçimlendirme dili olarak adlandırılmaktadır. CSS, web tabanlı projelerin geliştirilme aşamasında kullanılan 3 temel dilden birisidir. CSS ile resim, buton, metin, arka plan vb. birçok yapının renk ayarlarını, boyutlarını, konumlandırma şekillerinin yapısının vb. birçok şeyin ayarlanmasını ve yapılmasını sağlamaktadır.

HTML, aslına bakıldığında vakit metinlerin biçimlendirme veya yapılandırma alanında birçok olanak sağlamaktadır. Fakat HTML 'in ilk çıktığı dönemlerde yani CSS kullanılmadığı zamanlarda tüm web tabanlı projeler birbirine benzemektedir tabiri caizse eski televizyonlar yani siyah beyaz televizyonlar gibi görüntü vermektedir. CSS 'in ortaya çıkması ile birlikte şu anda bulunan web tabanlı projelerden de görüldüğü üzere hepsi birbirinden farklıdır. Yani CSS 'in kullanılması ile renkli televizyona geçiş sağlanmıştır (<https://www.webprofu.com.>, Erişim Tarihi: 01 Mayıs 2019).

Aşağıdaki avantajlar düşünülerek tez kapsamında CSS kullanılmıştır. Bu avantajlar;

- i. Web üzerinde yapılan uygulamaların çabuk bir şekilde indeksleyebilmesi
- ii. Web üzerinde yapılan uygulamaların istenilen şekilde veya yapıda yapılabilirliğini sağlaması
- iii. Bir veya daha fazla olan HTML dosya türlerinin tek merkezden yönetebilmesi

9.6. MySQL

Veritabanı yani database, birbirleri arasında bağlantı bulunan verilerin belirli bir sistem ve düzen doğrultusunda depolandığı yer olarak adlandırılmaktadır.

SQL, birbirleri ile ilişki içerisinde bulunan verilerin bulunduğu veri tabanında tutulan veriler üzerinde tasarım ve yönlendirme işlemlerinin yapılması için kullanılan bir sorgulama dili olarak adlandırılmaktadır. SQL birçok kişi tarafından programlama dili olarak bilinmesini rağmen aslında SQL içerisinde if, while gibi şart komutları barındırmadığından dolayı programlama dili değil, sorgulama dilidir.

MySQL, veri tabanındaki verileri yönetmek için kullanılan birden fazla SQL yöntemlerinden sadece bir tanesi olarak adlandırılmaktadır. MySQL, açık kaynak

kodlu veri tabanı sorgulama dili olarak adlandırılmaktadır. Web tabanlı uygulamalar geliştirilirken genellikle hızlı işlem yapması ve web tabanlı uygulama geliştirmek için Php dili tercih ediliyorsa bu dille en iyi şekilde çalışan veri tabanı yöntem bilimi olarak bilinmektedir. MySQL 'in tarihine bakıldığı vakit ilk olarak 1995 yılında piyasa çıkmış olduğu bilinmektedir (<https://www.mediatick.com.tr>., Erişim Tarihi: 09 Mart 2019)

Aşağıdaki avantajlar düşünülerek tez kapsamında MySQL kullanılmıştır. Bu avantajlar aşağıdaki gibidir;

- i. Windows, OS/2, UNIX gibi tüm işletim sistemlerinde çalışılabilir olması
- ii. Açık kaynak kodlu olması hem geliştirilebilmesini hem de öğrenilebilirliği kolaylaştırması
- iii. Kolay bir şekilde sorgulama yapısının bulunması
- iv. MySQL yazmak herhangi bir lisans ücretinin ödenmesinin gerekli olmaması
- v. MySQL 'in ön bellekleme bulundurması sayesinde daha evvelki zamanda zaman da yapılmış sorgulamaları kaydederek bir sonraki vakitte kullanılmasını kolaylaştırması
- vi. Tüm dünya alfabelerini ve dillerini depolayabilme yetisinin bulunması vb.

10. VERİ SETİNİN HAZIRLANMASI

Veri setinin hazırlanması başlığı altında geliştirilen tez de kullanılan veri setinin nitelik ve sınıf değerleri ile birlikte uygun veri sitenin oluşturulmasını ifade edilmiştir. Bununla birlikte sonuç ekranında ki bazı görseller anlatılmıştır.

Tezde kullanılan veri setini UCI Data Set sitesi üzerinden iki veri setinin birleşimi şeklinde oluşturulmuştur. Fakat bu iki veri seti yetersiz olduğundan bazı nitelik değerlerini banka danışmanları aracılığıyla elde edilip tez kapsamına dahil edildi. UCI Data Set üzerinden alınan veri setlerinden biri Portekiz bankasıyla alakalı olup diğeri Tayvan bankasıyla ilgilidir (<https://archive.ics.uci.edu/>, Erişim Tarihi: 09 Nisan 2019). Bu veri setlerinden esinlenerek ve danışmanların ışığında tez için daha etkin veri seti oluşturulması sağlandı ve hazırlanan veri seti GitHub 'a yüklenmiştir. Yüklenmiş olan veri setine belirtilmiş olan linkten ulaşılabilir ([https://github.com.](https://github.com/), Erişim Tarihi: 14 Temmuz 2019).

Tez çalışmasında, dolandırıcılığı tespit etmek için müşteriye ait eğitim verisinde on iki adet nitelik ve sınıf değeri mevcuttur. Her niteliğin içerisinde etiket (label) kolonunda iki adet değer vardır. Burada ki sayısı (count) kolonu veri setinde ilgili kategori içerisinde ki miktarını belirtir. Ağırlık (weight) kolonu ise kategorilerin boyutunu simgeler. Niteliklerin açıklanması aşağıdadır;

Kart Limiti: Müşterinin sahip olduğu kredi kart üzerinde ki harcayabileceği toplam miktardır. Müşterinin kart limit miktarı yedi grupta toplanmıştır. Aşağıdaki Şekil 10.1. 'de kart limitinin gruplanarak aralıklar halinde gösterimi bulunmaktadır.

Name: kartLimit Missing: 0 (0%)		Distinct: 6	Type: Nominal Unique: 0 (0%)
No.	Label	Count	Weight
1	500-1000	0	0.0
2	1000-3000	66	66.0
3	3000-6000	96	96.0
4	6000-10000	42	42.0
5	10000-20000	51	51.0
6	20000-30000	39	39.0
7	30000+	3	3.0

Şekil 10.1. Kart limitinin gruplandırılmış gösterimi

Maaş: Müşterinin aylık gelirinin gösterildiği niteliklerdir. Altı adet kategorik değer ile ifade edilmiştir. Aşağıdaki Şekil 10.2. 'de maaş niteliğinin gruplarla gösterimi bulunmaktadır.

Name: maas Missing: 0 (0%)		Distinct: 6	Type: Nominal Unique: 0 (0%)
No.	Label	Count	Weight
1	1000-3000	36	36.0
2	3000-6000	102	102.0
3	6000-10000	51	51.0
4	10000-20000	57	57.0
5	20000-50000	39	39.0
6	50000+	12	12.0

Şekil 10.2. Maaş niteliğinin gruplandırılmış gösterimi

Varlık Miktarı: Müşterinin toplam mal varlığının gösterildiği niteliklerdir. Yedi adet grup ile ifade edilmiştir. Aşağıdaki Şekil 10.3. 'de varlık miktarı niteliğinin gruplandırılmış gösterimi bulunmaktadır.

Name: varlikMiktari Missing: 0 (0%)		Distinct: 7	Type: Nominal Unique: 0 (0%)
No.	Label	Count	Weight
1	1000-10000	39	39.0
2	10000-30000	15	15.0
3	30000-60000	33	33.0
4	60000-100000	48	48.0
5	100000-300000	60	60.0
6	300000-500000	48	48.0
7	500000+	54	54.0

Şekil 10.3. Varlık miktarı niteliğinin gruplandırılmış gösterimi

Alışveriş Miktar Ortalaması: Müşterinin aylık olarak harcadığı ortalama alışveriş miktarıdır. Altı adet grup ile gösterilmiştir. Aşağıdaki Şekil 10.4. 'de alışveriş miktar ortalama niteliğinin gruplandırılmış gösterimi bulunmaktadır.

Name: alisverisMiktarOrt Missing: 0 (0%)		Distinct: 6	Type: Nominal Unique: 0 (0%)
No.	Label	Count	Weight
1	100-1000	81	81.0
2	1000-3000	84	84.0
3	3000-6000	60	60.0
4	6000-10000	24	24.0
5	10000-20000	45	45.0
6	20000-30000	3	3.0
7	30000+	0	0.0

Şekil 10.4. Alışveriş miktar ortalama niteliğinin gruplandırılmış gösterimi

Alışveriş Yiyecek: Müşterinin beş ay boyunca ilgili türden alışveriş yapma durumunu gösterir. İki kategorik değer ile ifade edilmiştir. Aşağıdaki Şekil 10.5. 'de alışveriş yiyecek niteliğinin kategorik gösterimi bulunmaktadır.

Name: alisverisYiyecek		Type: Nominal	
Missing: 0 (0%)		Distinct: 1	Unique: 0 (0%)
No.	Label	Count	Weight
1	aldi	297	297.0
2	almadi	0	0.0

Şekil 10.5. Alışveriş yiyecek niteliğinin kategorik gösterimi

Alışveriş Giyim: Müşterinin beş ay boyunca ilgili türden alışveriş yapma durumunu gösterir. İki kategori ile ifade edilmiştir. Aşağıdaki Şekil 10.6. 'da alışveriş giyim niteliğinin kategorik gösterimi bulunmaktadır.

Name: alisverisGiyim		Type: Nominal	
Missing: 0 (0%)		Distinct: 2	Unique: 0 (0%)
No.	Label	Count	Weight
1	aldi	291	291.0
2	almadi	6	6.0

Şekil 10.6. Alışveriş giyim niteliğinin kategorik gösterimi

Alışveriş Kozmetik: Müşterinin beş ay boyunca ilgili türden alışveriş yapma durumunu gösterir. İki adet kategori başlığı altında toplanmıştır. Aşağıdaki Şekil 10.7. 'de alışveriş kozmetik niteliğinin kategorik gösterimi bulunmaktadır.

Name: alisverisKozmetik		Type: Nominal	
Missing: 0 (0%)		Distinct: 2	Unique: 0 (0%)
No.	Label	Count	Weight
1	aldi	273	273.0
2	almadi	24	24.0

Şekil 10.7. Alışveriş kozmetik niteliğinin kategorik gösterimi

Alışveriş Mücevher: Müşterinin beş ay boyunca ilgili türden alışveriş yapma durumunu gösterir. İki kategori altında ifade edilmiştir. Aşağıdaki Şekil 10.8. 'de alışveriş mücevher niteliğinin kategorik gösterimi bulunmaktadır.

Name: alisverisMucevher		Type: Nominal	
Missing: 0 (0%)		Distinct: 2	Unique: 0 (0%)
No.	Label	Count	Weight
1	aldi	208	208.0
2	almadi	89	89.0

Şekil 10.8. Alışveriş mücevher niteliğinin kategorik gösterimi

Alışveriş Eğitim: Müşterinin beş ay boyunca ilgili türden alışveriş yapma durumunu gösterir. İki kategori ile ifade edilmiştir. Aşağıdaki Şekil 10.9. 'da alışveriş eğitim niteliğinin kategorik gösterimi bulunmaktadır.

Name: alisverisEgitim		Type: Nominal	
Missing: 0 (0%)		Distinct: 2	
		Unique: 0 (0%)	
No.	Label	Count	Weight
1	aldi	196	196.0
2	almadi	101	101.0

Şekil 10.9. Alışveriş eğitim niteliğinin kategorik gösterimi

Alışveriş Diğer: Müşterinin beş ay boyunca ilgili türden alışveriş yapma durumunu gösterir. İki kategori ile ifade edilmiştir. Aşağıdaki Şekil 10.10. 'da alışveriş diğer niteliğinin kategorik gösterimi bulunmaktadır.

Name: alisverisDiger		Type: Nominal	
Missing: 0 (0%)		Distinct: 2	
		Unique: 0 (0%)	
No.	Label	Count	Weight
1	aldi	220	220.0
2	almadi	77	77.0

Şekil 10.10. Alışveriş diğer niteliğinin kategorik gösterimi

Girdi Alışveriş Çeşit: Müşterinin yeni eklenecek alışverişinin türünü ifade eder. Altı kategori altında incelenmiştir. Aşağıdaki Şekil 10.11. 'de girdi alışveriş çeşit niteliğinin kategorik gösterimi bulunmaktadır.

Name: girdiAlisverisCesit		Type: Nominal	
Missing: 0 (0%)		Distinct: 6	
		Unique: 0 (0%)	
No.	Label	Count	Weight
1	yiyecek	27	27.0
2	giyim	45	45.0
3	kozmetik	66	66.0
4	mucevher	60	60.0
5	egitim	45	45.0
6	diger	54	54.0

Şekil 10.11. Girdi alışveriş çeşit niteliğinin kategorik gösterimi

Girdi Alışveriş Miktar: Müşterinin yeni yapacağı alışverişinin miktarını gösterir. Yedi adet grup altında incelenmiştir. Aşağıdaki Şekil 10.12. 'de girdi alışveriş miktar niteliğinin gruplandırılmış gösterimi bulunmaktadır.

Name: girdiAlisverisMiktar		Type: Nominal	
Missing: 0 (0%)		Distinct: 7	
		Unique: 0 (0%)	
No.	Label	Count	Weight
1	100-1000	78	78.0
2	1000-3000	57	57.0
3	3000-6000	72	72.0
4	6000-10000	33	33.0
5	10000-20000	45	45.0
6	20000-30000	9	9.0
7	30000+	3	3.0

Şekil 10.12. Girdi alışveriş miktar niteliğinin gruplandırılmış gösterimi

Yukarıda ki nitelikler işleme girerek sınıf değerini oluşturur. Sınıf değerleri az, orta, yüksek olmakla beraber üç adettir. Sınıfların veri setinde dağılımı aşağıdaki Şekil 10.13. 'de verildiği gibidir.

Name: riskGrubu		Type: Nominal	
Missing: 0 (0%)		Distinct: 3	
		Unique: 0 (0%)	
No.	Label	Count	Weight
1	az	75	75.0
2	orta	168	168.0
3	yuksekk	54	54.0

Şekil 10.13. Risk değerlerinin gösterimi

Müşteri için sınıflandırma işleminde kullanılan eğitim veri setinin, girdi nitelikleri ile birlikte verilerden birkaç tanesi aşağıdaki Şekil 10.14 'de gösterilmektedir.

kartLimit	maas	varlikMiktari	alisverisMiktarOrt	alisverisYiyecek	alisverisGiyim	alisverisKozmetik	alisverisMucevher	alisverisEgitim	alisverisDiger	girdiAlisverisCesit	girdiAlisverisMiktar	riskGrubu
3000-6000	3000-6000	60000-100000	100-1000	aldi	aldi	aldi	almadi	almadi	aldi	yiyecek	100-1000	az
1000-3000	6000-10000	30000-60000	100-1000	aldi	aldi	aldi	almadi	aldi	almadi	egitim	100-1000	az
3000-6000	6000-10000	500000+	100-1000	aldi	aldi	aldi	aldi	aldi	aldi	kozmetik	100-1000	az
1000-3000	3000-6000	10000-30000	100-1000	aldi	aldi	aldi	almadi	aldi	almadi	giyim	100-1000	az
1000-3000	3000-6000	300000-500000	100-1000	aldi	aldi	aldi	almadi	almadi	aldi	diger	100-1000	az
1000-3000	3000-6000	300000-500000	100-1000	aldi	aldi	aldi	aldi	almadi	aldi	mucevher	100-1000	az
3000-6000	3000-6000	60000-100000	100-1000	aldi	aldi	aldi	almadi	almadi	aldi	yiyecek	100-1000	az
1000-3000	3000-6000	60000-100000	100-1000	aldi	almadi	aldi	almadi	aldi	aldi	egitim	100-1000	az
1000-3000	3000-6000	300000-500000	100-1000	aldi	aldi	aldi	almadi	aldi	aldi	giyim	100-1000	az

Şekil 10.14. Müşteri eğitim veri seti gösterimi

Müşteri için oluşturulan ağaç modelinin sınanmasında kullanılan test veri setinin, girdi nitelikleri ile birlikte verilerden birkaç tanesi aşağıdaki Şekil 10.15 'de gösterilmektedir.

kartLimit	maas	varlikMiktari	alisverisMikt	alisverisYiyec	alisverisGiyir	alisverisKozn	alisverisMuc	alisverisEgiti	alisverisDige	girdiAlisveris	girdiAlisveris	riskGrubu
3000-6000	3000-6000	60000-10000	100-1000	aldi	aldi	aldi	almadi	almadi	aldi	yiyecek	100-1000	az
1000-3000	6000-10000	30000-60000	100-1000	aldi	aldi	aldi	almadi	aldi	almadi	egitim	100-1000	az
3000-6000	6000-10000	500000+	100-1000	aldi	aldi	aldi	aldi	aldi	aldi	kozmetik	100-1000	az
1000-3000	3000-6000	10000-30000	100-1000	aldi	aldi	aldi	almadi	aldi	almadi	giyim	100-1000	az
1000-3000	3000-6000	300000-50000	100-1000	aldi	aldi	aldi	almadi	almadi	aldi	diger	100-1000	az
1000-3000	3000-6000	300000-50000	100-1000	aldi	aldi	aldi	aldi	almadi	aldi	mucevher	100-1000	az
6000-10000	10000-20000	500000+	3000-6000	aldi	aldi	aldi	aldi	aldi	aldi	mucevher	3000-6000	az
6000-10000	6000-10000	60000-10000	1000-3000	aldi	aldi	aldi	aldi	aldi	aldi	egitim	1000-3000	az
3000-6000	6000-10000	60000-10000	100-1000	aldi	aldi	aldi	almadi	aldi	aldi	kozmetik	100-1000	az

Şekil 10.15. Müşteri test veri seti gösterimi

Dolandırıcılık risk değerlerinin belirlenmesinde, müşterinin beş aylık alışveriş alışkanlığı ve beş ay boyunca harcamış olduğu alışveriş miktarı nitelikleri en fazla öneme sahip değerlerdir. Bununla birlikte maaş, kart limit, varlık miktarı değerlerinin risk grubunu belirlemedeki etkileri daha azdır. Risk değerinin yüksek olması için müşterinin alışveriş alışkanlığına ve harcamasına uygun olamaması durumunu ifade etmektedir. Orta risk değerinde, alışveriş alışkanlığı veya alışveriş miktarı ortalama değerlerinden birini sağlamaması durumunu ifade eder. Az risk değerinde, müşteri alışveriş alışkanlıklarına uygun olarak satınalma işlemini gerçekleştirme durumunu ifade eder.

11. BULGULAR

Burada tez kapsamında kullanılan sınıflandırma yöntemi, karar ağacı sınıflandırma tekniğinin C4.5 algoritması kullanılarak ilgili veri seti üzerindeki ağaç yapısı ve ağaç içi kurallarının görselleri gösterilmiştir. Ayrıca ağaç yapısı üzerinde ki bazı kurallar “if- then” şeklinde sunumu sağlandı. Tez kapsamında on iki nitelik değeri ve buna bağlı sınıf değeri mevcuttur. Nitelik değerleri vasıtası ile dolandırıcılığın tespit edilmesi işlemi gerçekleştirilmiştir. Sınıflandırma işleminde kullanılacak sınıf değerleri üç kategori olarak ifade edilmiştir.

Bu kategoriler az, orta ve yüksek olmakla beraber dolandırma risk gruplarını belirler. Hiç bir zaman müşterinin dolandırılmadığının kesin olamamasından ötürü risksiz diye bir grup bulunmamaktadır. Niteliklere göre risk gruplarının belirlenmesi işleminde bir bilir kişiye danışılmıştır (Banka Danışmanı). Bilir kişiden alınan bilgiler ışığında satın alma geçmişinde olmayan ürünü alma ve aylık ortalamasının aşırı üstünde gibi durumlar göz önünde bulundurularak risk grupları belirlendi. İşlemlerin tamamı Weka programı üzerinde gerçekleştirildi. Eğitim verisinin Weka üzerine yüklenmesi sonucu oluşan karar ağacı Şekil 11.1. 'de ki gibidir.

Weka ile bireysel müşteri için oluşan ağaç modeline göre, girdi alışveriş miktarı 3000-6000 aralığında olanların şematik gösterimi aşağıdaki Şekil 11.3. 'de bulunmaktadır.

```
girdiAlisverisMiktar = 3000-6000
|
|   alisverisMiktarOrt = 100-1000: orta (0.0)
|   alisverisMiktarOrt = 1000-3000
|   |
|   |   girdiAlisverisCesit = yiyecek: orta (3.0)
|   |   girdiAlisverisCesit = giyim
|   |   |
|   |   |   varlikMiktari = 1000-10000: yuksek (0.0)
|   |   |   varlikMiktari = 10000-30000: orta (3.0)
|   |   |   varlikMiktari = 30000-60000: yuksek (3.0)
|   |   |   varlikMiktari = 60000-100000: yuksek (3.0)
|   |   |   varlikMiktari = 100000-300000: yuksek (0.0)
|   |   |   varlikMiktari = 300000-500000: yuksek (0.0)
|   |   |   varlikMiktari = 500000+: yuksek (0.0)
|   |   |
|   |   |   girdiAlisverisCesit = kozmetik
|   |   |   |
|   |   |   |   varlikMiktari = 1000-10000: yuksek (0.0)
|   |   |   |   varlikMiktari = 10000-30000: yuksek (0.0)
|   |   |   |   varlikMiktari = 30000-60000: yuksek (0.0)
|   |   |   |   varlikMiktari = 60000-100000: orta (3.0)
|   |   |   |   varlikMiktari = 100000-300000: yuksek (3.0)
|   |   |   |   varlikMiktari = 300000-500000: yuksek (3.0)
|   |   |   |   varlikMiktari = 500000+: yuksek (0.0)
|   |   |   |
|   |   |   |   girdiAlisverisCesit = mucevher: yuksek (6.0)
|   |   |   |   girdiAlisverisCesit = egitim: yuksek (0.0)
|   |   |   |   girdiAlisverisCesit = diger: orta (3.0)
```

Şekil 11.3. Girdi alışveriş miktarı 3000-6000 kurallarının gösterimi

Yukarıda bulunan Şekil 11.3. 'de şematik olarak gösterimi yapılmakta olan kurallardan bazılarının eğer – ise (if - then) biçiminde gösterilmektedir. Bunlar;

- 1) Eğer $girdiAlisverisMiktar = 1000-3000$ ve $kartLimit = 1000 - 3000$ ve $varlikMiktari = 1000-10000$ ve $alisverisDiger=almadi$ ve $alisverisMucevher = aldi$ Then $riskGrubu = yüksek$
- 2) Eğer $girdiAlisverisMiktar = 1000 - 3000$ ve $kartLimit = 1000-3000$ ve $varlikMiktari = 1000-10000$ ve $alisverisDiger=almadi$ ve $alisverisMucevher = almadi$ ve $alisverisKozmetik = aldi$ Then $riskGrubu = orta$
- 3) Eğer $girdiAlisverisMiktar = 1000-3000$ ve $kartLimit = 1000-3000$ ve $varlikMiktari = 1000-10000$ ve $alisverisDiger=almadi$ ve $alisverisMucevher = almadi$ ve $alisverisKozmetik = almadi$ Then $riskGrubu = yüksek$

Weka ile bireysel müşteri için oluşan ağaç modeline göre, girdi alışveriş miktarı 6000-10000 aralığında olanların şematik gösterimi aşağıdaki Şekil 11.4. 'de bulunmaktadır.

```

| alisverisMiktarOrt = 3000-6000
|   maas = 1000-3000: orta (0.0)
|   maas = 3000-6000: orta (3.0)
|   maas = 6000-10000: orta (12.0)
|   maas = 10000-20000
|     varlikMiktari = 1000-10000: az (0.0)
|     varlikMiktari = 10000-30000: az (0.0)
|     varlikMiktari = 30000-60000: az (0.0)
|     varlikMiktari = 60000-100000: az (0.0)
|     varlikMiktari = 100000-300000: orta (3.0)
|     varlikMiktari = 300000-500000: orta (9.0/3.0)
|     varlikMiktari = 500000+: az (6.0)
|   maas = 20000-50000: az (3.0)
|   maas = 50000+: orta (0.0)
| alisverisMiktarOrt = 6000-10000: orta (6.0)
| alisverisMiktarOrt = 10000-20000: orta (0.0)
| alisverisMiktarOrt = 20000-30000: orta (0.0)
| alisverisMiktarOrt = 30000+: orta (0.0)
| girdiAlisverisMiktar = 6000-10000
|   varlikMiktari = 1000-10000: orta (0.0)
|   varlikMiktari = 10000-30000: orta (0.0)
|   varlikMiktari = 30000-60000: yuksek (3.0)
|   varlikMiktari = 60000-100000
|     maas = 1000-3000: orta (0.0)
|     maas = 3000-6000: yuksek (3.0)
|     maas = 6000-10000: orta (6.0)
|     maas = 10000-20000: orta (0.0)
|     maas = 20000-50000: orta (0.0)
|     maas = 50000+: orta (0.0)
|   varlikMiktari = 100000-300000: yuksek (3.0)
|   varlikMiktari = 300000-500000: orta (3.0)
|   varlikMiktari = 500000+: orta (15.0)

```

Şekil 11.4. Girdi alışveriş miktarı 6000-10000 kurallarının gösterimi

Yukarıda bulunan Şekil 11.4. 'de şematik olarak gösterimi yapılmakta olan kurallardan bazılarının eğer – ise (if - then) biçiminde gösterilmektedir. Bunlar;

- 1) Eğer $girdiAlisverisMiktar = 3000-6000$ ve $alisverisMiktarOrt = 1000-3000$ ve $girdiAlisverisCesit = yiyecek$ Then $riskGrubu = orta$
- 2) Eğer $girdiAlisverisMiktar = 3000-6000$ ve $alisverisMiktarOrt = 1000-3000$ ve $girdiAlisverisCesit = giyim$ ve $varlikMiktari = 10000-30000$ Then $riskGrubu = orta$
- 3) Eğer $girdiAlisverisMiktar = 3000-6000$ ve $alisverisMiktarOrt = 1000-3000$ ve $girdiAlisverisCesit=giyim$ ve $varlikMiktari = 60000-100000$ Then $riskGrubu = yuksek$

Weka ile bireysel müşteri için oluşan ağaç modeline göre, girdi alışveriş miktarı 10000-20000 ve 20000-30000 aralığında olanların şematik gösterimi aşağıdaki Şekil 11.5. 'de bulunmaktadır.

```

girdiAlisverisMiktar = 10000-20000
|   alisverisMiktarOrt = 100-1000: orta (0.0)
|   alisverisMiktarOrt = 1000-3000: orta (0.0)
|   alisverisMiktarOrt = 3000-6000: yuksek (6.0)
|   alisverisMiktarOrt = 6000-10000
|   |   varlikMiktari = 1000-10000: yuksek (0.0)
|   |   varlikMiktari = 10000-30000: yuksek (0.0)
|   |   varlikMiktari = 30000-60000: orta (3.0)
|   |   varlikMiktari = 60000-100000: yuksek (0.0)
|   |   varlikMiktari = 100000-300000: yuksek (0.0)
|   |   varlikMiktari = 300000-500000: yuksek (3.0)
|   |   varlikMiktari = 500000+: yuksek (3.0)
|   alisverisMiktarOrt = 10000-20000: orta (30.0)
|   alisverisMiktarOrt = 20000-30000: orta (0.0)
|   alisverisMiktarOrt = 30000+: orta (0.0)
girdiAlisverisMiktar = 20000-30000
|   girdiAlisverisCesit = yiyecek: orta (0.0)
|   girdiAlisverisCesit = giyim: orta (0.0)
|   girdiAlisverisCesit = kozmetik: orta (0.0)
|   girdiAlisverisCesit = mucevher: orta (3.0)
|   girdiAlisverisCesit = egitim: orta (3.0)
|   girdiAlisverisCesit = diger: yuksek (3.0)
girdiAlisverisMiktar = 30000+: orta (3.0)

```

Şekil 11.5. Girdi alışveriş miktarı 10000-20000 ve 20000-30000 kurallarının gösterimi

Yukarıda bulunan Şekil 11.5. 'de şematik olarak gösterimi yapılmakta olan kurallardan bazılarının eğer – ise (if - then) biçiminde gösterilmektedir. Bunlar;

- 1) Eğer $girdiAlisverisMiktar = 3000-6000$ ve $alisverisMiktarOrt = 1000-3000$ ve $girdiAlisverisCesit = mücevher$ Then $riskGrubu = yüksek$
- 2) Eğer $girdiAlisverisMiktar = 3000-6000$ ve $alisverisMiktarOrt=1000-3000$ ve $girdiAlisverisCesit = diğer$ Then $riskGrubu = orta$
- 3) Eğer $girdiAlisverisMiktar = 3000-6000$ ve $alisverisMiktarOrt = 1000-3000$ ve $girdiAlisverisCesit = kozmetik$ ve $varlikMiktari = 300000 - 500000$ Then $riskGrubu= yüksek$

Sınıfı ve nitelikleri belirlenip oluşturulan veri setini Weka üzerindeki sınıflandırma modeli ve J48 (C4.5 algoritması ile) işleme soktuğumuzda karar ağacı tekniğine uygun olarak işleme alındığında aşağıdaki Şekil 11.6. 'da gösterilen sonuçlar alınmıştır.

=== Summary ===

Correctly Classified Instances	282	94.9495 %
Kappa statistic	0.9148	
Mean absolute error	0.0531	
Root mean squared error	0.1629	
Relative absolute error	13.6264 %	
Root relative squared error	36.9401 %	
Total Number of Instances	297	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.960	0.054	0.857	0.960	0.906	0.874	0.978	0.889	az
	0.929	0.023	0.981	0.929	0.954	0.900	0.983	0.981	orta
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	yukse
Weighted Avg.	0.949	0.027	0.953	0.949	0.950	0.911	0.985	0.961	

=== Confusion Matrix ===

a	b	c	<-- classified as
72	3	0	a = az
12	156	0	b = orta
0	0	54	c = yuksek

Şekil 11.6. Hata oranlarının gösterimi

Yukarıdaki Şekil 11.6. 'da Weka yazılımı aracılığı ile oluşan ağacın modeli sonucunda ortaya çıkan hata oranlarının gösterimi bulunmaktadır. Burada karmaşıklık matrisi (confusion matrix) başlığı altında gösterilen tablo yorumlanacak olunursa; a, az riski, b, orta riski, c, yüksek riski ifade etmektedir. Az riskte 75 adet müşteri, orta riskte 168 adet müşteri, yüksek riskte 54 adet müşteri bulunmaktadır. Şekil 11.6. 'da karmaşıklık matrisi üzerindeki diagonal çizgi doğru sınıflandırma ifade eder. Diagonal çizginin dışında kalanlar ise yanlış sınıflandırmayı ifade etmektedir.

Elde edilen risk değerlerine göre tez çalışmasında yapılacak olan işlemler aşağıdaki Şekil 11.7. 'de gösterilmektedir.

SINIF DEĞERİ	AÇIKLAMA
AZ	Satınalma işleminin gerçekleştirilmesini sağlar.
ORTA	Müşteriye SMS yolu ile satınalma işlemi gerçekleştirilmesi sağlanır.
YUKSEK	Müşterinin kredi kartının aktiflik durumu inaktife çekilerek satınalma işlemi engellenir.

Şekil 11.7. Risk değerlerine göre yapılacak işlemler



12. UYGULAMA ARAYÜZLERİ

Planı yapıp uygulaması gerçekleştirilen tezin uygulama arayüz görselleri aşağıdaki başlıklar altında gösterilmektedir. Tezin uygulamasının arayüz tasarımını yaparken CSS, HTML, PHP teknolojilerinden yardım alınmıştır.

12.1. Uygulama Giriş Sayfası Arayüzü

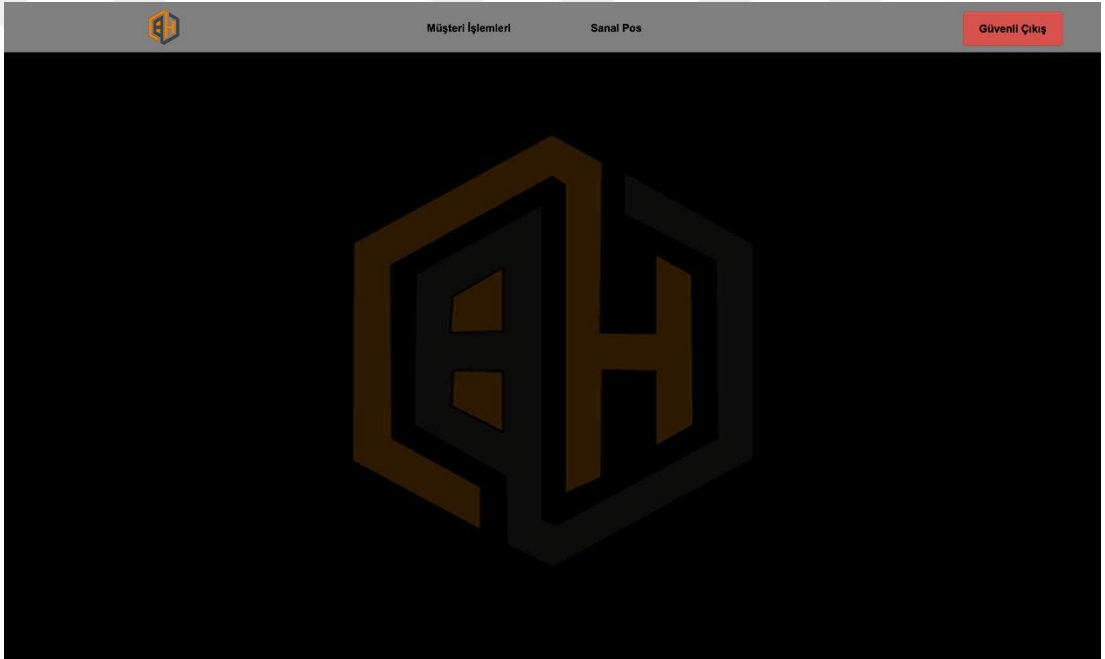
Uygulamanın giriş sayfası arayüzünde, kullanıcı girişi işlemlerini sağlayan sayfa bulunmaktadır. Bu sayfada kullanıcı adı ve şifresi girdi olarak alınarak doğruluğu kontrol edildikten sonra uygulamanın ana sayfa arayüzüne geçiş işleminin yapılmasını gerçekleştirmektedir. Tasarımı yapılmış olan uygulamanın kullanıcı giriş sayfası ekranı aşağıda Şekil 12.1. 'de gösterilmektedir.



Şekil 12.1. Uygulama kullanıcı giriş ekranı

12.2. Uygulama Anasayfa Arayüzü

Tez uygulamasının anasayfa arayüzü, projenin belirtilen koşulları yerine getirebilmesi için yönlendirilmelerin bulunduğu sayfadır. Kullanıcı bu sayfa üzerinde bulunan müşteri işlemleri ve sanal pos işlemleri yazılarına basarak yapmak istediği işlemleri seçebilir. Müşteri işlemleri yazısına bastığında müşteri ilgili işlemlerin yapılacağı sayfaya, sanal pos yazısına basarak alışveriş girdilerinin alındığı ve dolandırıcılığın tespit edilmesi işleminin yapıldığı sayfalara yönlendirilmesi sağlanmaktadır. Anasayfada sağ üst köşede bulunan buton yardımı ile uygulamadan çıkış sağlanmaktadır. Tasarımı yapılmış olan uygulamanın anasayfa ekranı aşağıda Şekil 12.2. 'de gösterilmektedir.



Şekil 12.2. Uygulama anasayfa ekranı

12.3. Uygulama Sanal Pos Arayüzü

Tez uygulamasının sanal pos arayüzünde, bankaya ait olan müşterilerin kredi kartı ile herhangi bir mağazadan alışveriş yapabilmesini sağlayan sayfadır. Burada müşteri, gerekli kayıtları doldurarak alışveriş eylemini gerçekleştirir. Alışveriş Ekle butonuna basıldığı takdirde Risk grubu sayfasına yönlendirilerek müşterinin risk grubu ve yapılacak işlem kullanıcıya sunulur. Tasarımı yapılmış olan uygulamanın sanal pos işlemleri ekranı aşağıda Şekil 12.3. 'de gösterilmektedir.

Label	Input Field
Kart Limit	
Maas	
Varlık Miktarı	
Alışveriş Miktar Ort	
Alışveriş Yıyacak	
Alışveriş Giyim	
Alışveriş Kozmetik	
Alışveriş Mücevher	
Alışveriş Eğlim	
Alışveriş Diğer	
Girdi Alışveriş Çeşit	
Girdi Alışveriş Miktar	

Şekil 12.3. Uygulama sanal pos ekranı

12.4. Uygulama Müşteri Risk Grubu Arayüzü

Tez uygulamasının müşteri risk grubu arayüzünde, bankaya ait olan müşterilerin alışveriş alışkanlıkları dışında bir işlem seyir etmesi durumunda müşterinin risk grubunun ve alışveriş gerçekleştirme durumunun gösterildiği sayfadır. Tasarımı yapılmış olan uygulamanın müşteri risk grubu ekranı aşağıda Şekil 12.4. 'de gösterilmektedir.

Label	Input Field
Risk Grubu	az
Alışveriş Gerçekleştirme Durumu	İşlem Gerçekleştiriliyor...

Şekil 12.4. Uygulama müşteri risk grubu ekranı

12.5. Uygulama Müşteri İşlemleri Arayüzü

Tez uygulamasının müşteri işlemleri arayüzünde, bankaya ait olan müşteriler üzerinde işlemlerin yapılabilmesi sağlanmaktadır. Burada kullanıcı müşteri ekle,

müşteri sil, müşteri güncelle yazılarından istenilen işlemleri yapmak için bastığında ilgili sayfalara yönlendirilerek yapılması istenen sayfanın arayüzü ile karşılaşmasını sağlamaktadır. Ayrıca bu sayfa üzerinde müşteri listelerinin görüntülenebileceği tablo bulunmaktadır. Tasarımı yapılmış olan uygulamanın müşteri işlemleri ekranı aşağıda Şekil 12.5. 'de gösterilmektedir.

KART NO	MÜŞTERİ ADI	MÜŞTERİ SOYADI	CINSİYET	YAŞ	MAAŞ	KART LİMİT	VARLIK MİKTARI	ÖĞRENİM	MESLEK	ADRES	TELEFON	MAİL
101111	Orkun	İsi7k	erkek	39	21000	20000	60000	yuksekoğretim	muhendis	Beylikd7uzu?	2147483647	orkunisi7k@gmail.com
110111	?Saban	K7ok	erkek	36	8000	4000	140000	yuksekoğretim	muhendis	Tuzla	2147483647	sabankucuk@hotmail.com
111011	Ebru	Sak	kadin	28	5000	3000	170000	yuksekoğretim	ogretmen	Gu7ngoren	2147483647	ebrusak@gmail.com
111101	Serpil	Cimsir	kadin	45	3800	3000	85000	ortaoğretim	memur	Besiktas	2147483647	serpicimsir@outlook.com
111110	Emre	Sar7i	erkek	35	5500	15000	75000	yuksekoğretim	avukat	Zeytinburnu	2147483647	emre_sari53@outlook.com
111111	Ipek	Y7ilmaz	kadin	30	3000	2000	24000	ortaoğretim	memur	Atasehir	2147483647	ipek@hotmail.com
111112	P7inar	Birinci	kadin	39	12000	10000	50000	yuksekoğretim	doktor	Beyoglu	2147483647	p7inarbirinci@outlook.com
111113	Ece	Keles	kadin	25	5000	7000	75000	yuksekoğretim	ogretmen	Kad7ikoy	2147483647	ecakeles@gmail.com
111114	Ayse	Badem	kadin	38	5000	3000	110000	yuksekoğretim	avukat	Esenyurt	2147483647	aysebadem.@hotmail.com
111115	Esra	Keles	kadin	37	50000	45000	100000	ortaoğretim	diger	Kartal	2123205621	esrakeles@msn.com
111116	Cemil	Eksi	erkek	32	1890	3000	20000	ilkogretim	emekli	Avc7ilar	2147483647	cemileksi@hotmail.com
111117	Hatice	Sendir	kadin	32	4500	4000	50000	ilkogretim	diger	Cekmekoy	2147483647	haticesendir@outlook.com
111118	Eda	Ko7il	kadin	41	17000	7000	200000	yuksekoğretim	doktor	Umraniye	2147483647	edakotil@gmail.com
111119	H7ulya	Bulut	kadin	31	14000	10000	180000	yuksekoğretim	avukat	Sancaktepe	2147483647	h7ulyabulut@gmail.com
111121	Sena	Mese	kadin	25	2800	1000	6000	yuksekoğretim	memur	Sultanbeyli	2147483647	senamese@outlook.com
111131	Sevda	Akp7inar	kadin	36	6500	5000	70000	yuksekoğretim	ogretmen	Adalar	2147483647	sevdaakp7inar@gmail.com
111141	?Song7ul	Nane	kadin	36	4000	3000	240000	yuksekoğretim	memur	Arnavutkoy	2147483647	sengulnane@gmail.com

Şekil 12.5. Uygulama müşteri işlemleri ekranı

12.6. Uygulama Müşteri Ekle Arayüzü

Tez uygulamasının müşteri ekle arayüzünde, bankaya yeni kayıt olmak isteyen müşterilerin gerekli bilgileri girdi olarak istenilerek kayıt altında tutulması işlemini sağlayan sayfadır. Bu sayfada bulunan devam et butonuna tıklayarak müşteri bilgileri ile beraber veri tabanında tutulma işlemi gerçekleştirilir. Tasarımı yapılmış olan tez uygulamasının müşteri ekle ekranı aşağıda Şekil 12.6 'da gösterilmektedir.

KART NO	
MÜŞTERİ AD	
MÜŞTERİ SOYAD	
CİNSİYET	
YAŞ	
MAAŞ	
KART LIMIT	
VARLIK MİKTARI	
ÖĞRENİM	
MESLEK	
ADRES	
TELEFON	
MAIL	

Kayıt Ekle

Şekil 12.6. Uygulama müşteri ekle ekranı

12.7. Uygulama Müşteri Güncelle Arayüzü

Tez uygulamasının müşteri güncelle arayüzünde, bankada kaydı bulunan yani bankanın müşterisi olan kişilerin gerekli bilgileri üzerinde değişiklik yapılarak tekrardan kayıt edilmesi işlemini sağlayan sayfadır. Bu sayfada bulunan devam et butonuna tıklayarak müşterinin yeni bilgileri veri tabanında tutulması gerçekleştirilir. Tasarımı yapılmış olan uygulamanın müşteri güncelle ekranı aşağıda Şekil 12.7. 'de gösterilmektedir.

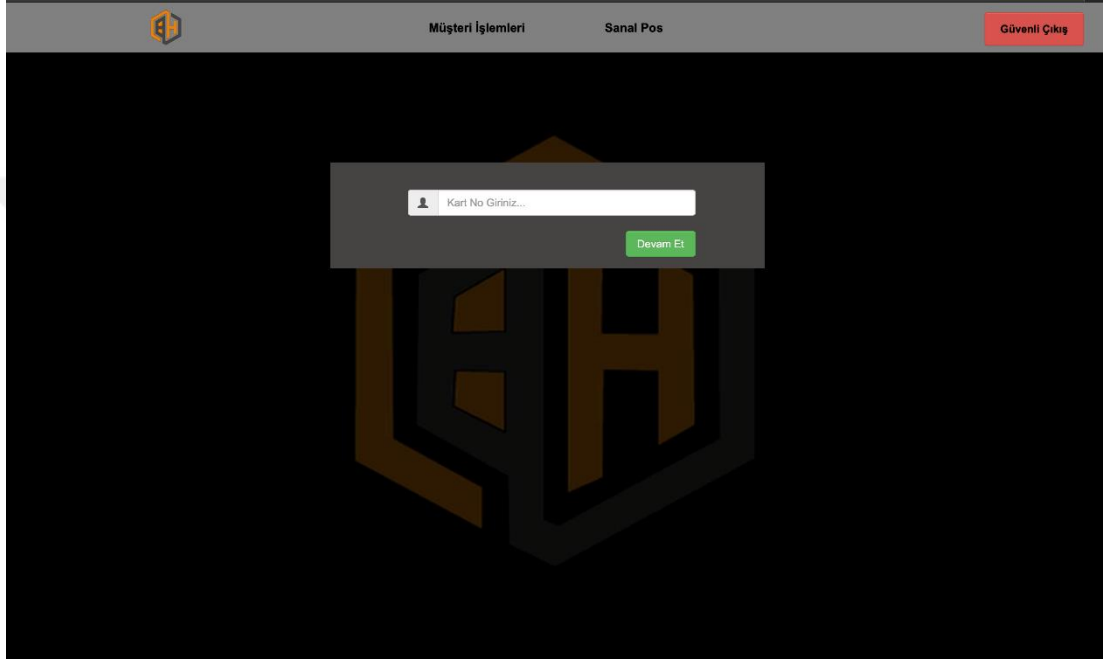
KART NO	
MÜŞTERİ AD	Orkun
MÜŞTERİ SOYAD	İsirtk
CİNSİYET	erkek
YAŞ	39
MAAŞ	21000
KART LIMIT	20000
VARLIK MİKTARI	60000
ÖĞRENİM	yuksekoğretim
MESLEK	muhendis
ADRES	Beylikdüzü?
TELEFON	2147483647
MAIL	orkuniskik@gmail.com

Güncelle

Şekil 12.7. Uygulama müşteri güncelle ekranı

12.8. Uygulama Müşteri Sil Arayüzü

Tez uygulamasının müşteri sil arayüzünde, bankada kaydı bulunan yani bankanın müşterisi olan kişinin bankadan ayrılmak istediği durumda gerekli girdi bilgisi alındıktan sonra silinmesi işlemini sağlayan sayfadır. Bu sayfada bulunan devam et butonuna tıklayarak müşterinin bilgileri veri tabanından silinme işlemi gerçekleştirilir. Tasarımı yapılmış olan uygulamanın müşteri sil ekranı aşağıda Şekil 12.8. 'de gösterilmektedir.



Şekil 12.8. Uygulama müşteri sil ekranı

12.9. Uygulama Alışveriş Listeleme Arayüzü

Tez uygulamasının alışveriş listeleme arayüzünde, bankada kaydı bulunan yani bankanın müşterisi olan kişinin geçmiş son beş aya ait alışveriş türlerinin ve miktarlarının listelendiği arayüz sayfasıdır. Tasarımı yapılmış olan uygulamanın alışveriş listeleme ekranı aşağıda Şekil 12.9. 'da gösterilmektedir.



ALIŞVERİŞ LİSTELEME

KART NO	MİKTAR AY1	MİKTAR AY2	MİKTAR AY3	MİKTAR AY4	MİKTAR AY5	ÇEŞİT AY1	ÇEŞİT AY2	ÇEŞİT AY3	ÇEŞİT AY4	
111111	1300	1200	750	1000	1100	yiyecek,kozmetik	kozmetik,giyim,yiyecek	yiyecek,diger	yiyecek,kozmetik	
111112	4500	3200	7000	5000	3000	yiyecek,egitim,giyim	egitim,kozmetik,yiyecek	egitim,kozmetik,yiyecek,giyim	yiyecek,egitim,giyim	e
111113	2500	4800	3000	3500	4800	yiyecek,giyim,kozmetik	kozmetik,mucevher,yiyecek	kozmetik,mucevher,yiyecek,giyim	yiyecek,giyim,kozmetik,diger	koz
111114	1200	1700	1800	1400	1600	giyim,yiyecek,egitim	giyim,kozmetik,egitim	giyim,yiyecek,egitim	giyim,diger,yiyecek	
111115	35000	37500	34250	30150	32100	egitim,diger,mucevher	diger,egitim,yiyecek,kozmetik	diger,egitim,mucevher	egitim,diger,diger,yiyecek	diger.e
111116	1400	1800	1500	1550	1650	giyim,yiyecek,kozmetik,diger	yiyecek,kozmetik,diger	yiyecek,kozmetik,egitim	giyim,yiyecek,egitim	yiye
111117	2800	2100	1900	1700	1850	diger,yiyecek,giyim	diger	egitim,yiyecek,giyim,yiyecek	mucevher,yiyecek,kozmetik	d
111118	4800	3800	4500	4350	4100	giyim,kozmetik,mucevher	yiyecek,kozmetik,egitim	yiyecek,kozmetik,giyim,egitim	giyim,kozmetik,mucevher	yiyece
111119	3000	3250	2900	3200	2850	yiyecek,giyim,egitim	kozmetik,yiyecek,egitim	kozmetik,yiyecek,egitim,giyim	yiyecek,giyim,egitim,diger	k
111110	1900	2100	2400	2500	1350	kozmetik,giyim,yiyecek,egitim	yiyecek,diger	yiyecek,giyim,kozmetik,egitim	mucevher,giyim,yiyecek	
111121	500	400	700	550	685	yiyecek	kozmetik,yiyecek	giyim,yiyecek	yiyecek,diger	yi
111131	2000	2500	3000	3700	3300	yiyecek,diger,egitim	giyim,kozmetik,egitim,yiyecek	yiyecek,giyim,mucevher	yiyecek,diger,egitim	giyim
111141	1800	1950	1780	1920	1975	yiyecek,kozmetik,giyim,diger	giyim,kozmetik,egitim	yiyecek,diger,egitim	yiyecek,diger,giyim	giyim
111151	1400	1650	2500	3000	1650	yiyecek,diger	kozmetik,giyim,diger	kozmetik,yiyecek,diger	yiyecek,diger	
111161	600	750	1000	850	975	giyim,kozmetik	yiyecek,giyim	yiyecek,diger,yiyecek	diger,yiyecek	

Şekil 12.9. Uygulama alışveriş listeleme ekranı

13. TARTIŞMA

Literatürde yer alan bankacılık sektöründe dolandırıcılık örneklerinin oluşturduğu katkılar karşılaştırılarak incelenmiştir. Elde edilen sonuçlar yorumlanmıştır. Kullanılmış veri setleri üzerinde bilgilerin gizliliği ve bankacılık sektöründeki dolandırıcılık çeşitleri baz alınarak kesin ve net bir karşılaştırmanın yapılması mümkün değildir. Literatürde internet bankacılığı, elektronik ticaret gibi yapımları bulunmakla beraber gerçekleştirilen tezde fiziksel mağazalar üzerinde yapılabilecek kredi kartı tabanlı dolandırıcılığın önüne geçme hedeflenmiştir.

Yapılan tezde, sınıflandırılma işlemi için kullanılan eğitim veri seti direkt olarak UCI dataset sitesinden alınmayıp, buradaki veri setlerinden bir eğitim veri seti oluşturulmuştur. Veri seti, tez konusu benzerlerine göre farklılık göstermekle beraber tahmin edici bir modeldir. Benzer yapılan tezlerde ise, direkt olarak UCI dataset sitesinde hazır olarak bulunan veri setleri üzerinde sınıflandırma işlemlerinin yapıldığı gözlemlenmiştir.

Literatürde bankacılık sektörü üzerindeki çeşitli veri setini birçok farklı veri madenciliği modelleri, teknikleri ve buna bağlı olarak çeşitli algoritmalar denenmiş ve farklı sonuçlar alınmıştır. Kullanılan modeller genelde kümeleme, sınıflandırma ve birliktelik kuralları tabanlı olup farklı yaklaşımlar da tespit edilmiştir. Aynı zamanda bu modeller üzerinde farklı teknik ve algoritmalar da uygulanmıştır. Genel başarı oranları değişmekle beraber %80 ve %90 arasındadır. Geliştirilen tez veri madenciliği sınıflandırma yöntemi üzerinde karar ağacı ile sınıflandırma tekniği ve C4.5 algoritması uygulanarak %10 test veri seti üzerinden %90 lık bir eğitimle %88 lik bir başarı sağlanmıştır.

Tez kapsamında on iki farklı nitelik ile beraber bir sınıflandırma işlemi gerçekleştirildi. Tez beş aylık müşteri satın alma alışkanlıklarını üzerine kurulu olmakla

beraber bu, müşterilerin yeni alışverişlerinde dolandırma ihtimallerinin önüne geçerek önleyici bir model olma yetisini kazanmıştır.

Literatürdeki benzer tezde (Soylu K. (2018) Kredi Kartı Sahte İşlem Tespiti, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara, Danışman: Prof.Dr. Şahin Emrah) veri madenciliği yapılması için farklı yazılımlardan destek alınmıştır. Bu yazılımlar;

- i. WEKA
- ii. SPSS
- iii. RapidMiner
- iv. R Programing Vb.

Yapılan tezde Weka yazılımı ile sınıflandırma işlemi yapıldı. Weka yazılımının seçilme nedeni kullanım kolaylığı ve esnekliği aynı zamanda tüm platformlarda hizmet edebilme özelliğidir.

Benzer yapılmış tezde (Soylu K. (2018) Kredi Kartı Sahte İşlem Tespiti, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara, Danışman: Prof.Dr. Şahin Emrah) genellikle veri madenciliği yazılımları üzerinden çalışma gösterilmiştir. Fakat gerçekleştirilen tezde Weka gibi veri madenciliği yazılımları amaç olarak değil araç olarak görülmüştür. Weka 'da yapılan sınıflandırma sonucu oluşan ağaç modeli sonucunun HTML, CSS ile oluşturulan PHP tabanlı bir web arayüze aktarılması sağlanmaktadır. Bununla beraber görsel olarak daha tatminkâr ve erişim kolaylığı gibi yetiler tez kapsamına kazandırılmıştır.

14. SONUÇ

Dünyanın küreselleşmesi ve genelleşmesi bazı ihtiyaçların önem derecesini arttırmıştır. İnsanların günlük yaşantılarında hız ve zaman kavramları çok önemli bir noktaya erişti. Bankacılık ve finans sektörü de bu kavramların önemlerini baz alarak internet ve sanal dünyaya girmek zorunda kalmıştır. Sanal dünyada, bankalara daha az iş gücü, maliyet ve daha fazla müşteri gibi birçok imkân sağlamaktadır. Lakin sanal dünyanın birtakım zararları da vardır. Oluşabilecek en büyük zarar güvenliktir. Çünkü sektörde ki paranın dolaşımı dolandırıcıların iştahını kabartmıştır. Bütün bu olasılıklar düşünüldüğü takdirde bankacılık ve finans sektörü de oluşabilecek dolandırıcılığı, tespit ve önleme konusunda çalışmaya karar verildi.

Geliştirilmiş tezde öncelikle bankacılık sektörü konusunda detaylı bir çalışma yapıldı, sonrasında dolandırıcılık kavramı nedir ve türleri nelerdir gibi konularda araştırmalar yapılarak iki temel konu birbirleriyle bağdaştırıldı. Aynı zamanda geliştirilmiş tezin ilgili teknolojileri ve veri madenciliği yöntemleri detaylı olarak incelendi. Tezin, sektöre ne gibi yararlar getireceği düşünüldü. Gerçekleştirilen kazanımlar sonucu tez tamamlanmış ve dolandırıcılığın tespit yetisi sisteme kazandırılmıştır.

Tezde öncelikle benzer çalışmalar analiz edilip konu hakkında kullanılan veri madenciliği modelleri incelendi. Araştırma sonucu en yüksek başarıyı verecek model olarak sınıflandırma yöntemi seçildi. Yöntemin, karar ağacı ile sınıflandırma tekniği ile kolay okunabilirlik ve analiz gibi avantajları kullanıldı. Tekniğin uygun algoritması olarak C4.5 algoritması seçildi. C4.5 algoritmasının kategorik veriler dışında sayısal verileri üzerinde de çalışabilmesi gibi özellikleri düşünülüp tezin geliştirimi için C4.5 algoritmasında karar kılındı. Veri madenciliğinin yapılması için Weka yazılımı seçildi.

Weka yazılımında sınıflandırma işleminin yapılabilmesi için gerekli eğitim veri seti oluşturuldu. Sınıflandırılma işlemi için kullanılan eğitim veri seti direkt olarak UCI dataset sitesinden alınmayıp, buradaki veri setlerinden (Tayvan ve Portekiz) esinlenerek sentetik bir eğitim veri seti oluşturulmuştur. Sentetik veri seti oluşturulurken banka danışmanına başvurulup danışmanın bilgi birikimi kullanılarak veri setinin hazırlanması tamamlanmıştır. Banka danışmanı aracılığıyla günümüz koşullarına en yakın veri seti elde edilmiştir. Bütün bu işlemler sonucu oluşturulan sentetik eğitim veri seti tez konusu benzerlerine göre farklılık göstermekle beraber tahmin edici bir modeldir.

On iki farklı nitelik değeri üzerinde çalışıldı. Weka üzerinde oluşturulan ağaç yapısından modelin boyutu 110, yaprak sayısı 91 adet ve modelin oluşturulma süresi 0.02 saniye olarak gözükmektedir. Oluşturulmuş karar ağacı üzerindeki 297 adet verinin 282 tanesi doğru sınıflandırılarak sınıf değerlerinin başarı oranı %94.9495 olarak ortaya çıkmıştır. Cross – validation oranı da (veri setinin %10' luk kısmı test edilerek kalan veri setinin %90' ında eğitim işlemi yapılacak) %88.88 olarak ortaya çıkmıştır. Bununla beraber Kappa Statistic 0.8107 ve ortalama mutlak hata değeri 0.0955' dir. Ayrıca 43 adet veriden oluşan bir test veri seti hazırlanarak oluşturulan karar ağacı üzerinde sınıflandırma işlemine sokulmuştur. Gerçekleşen işlem sonucu test veri setinde bulunan 43 adet veriden 41 tanesi doğru olarak sınıflandırılmaktadır. İşlem sonucu oluşan genel başarı Weka arayüzünün belirttiği değerler ile başarı oranı % 95.3488 olarak elde edilmiştir. Burada oluşturulan karar ağacının kuralları PHP üzerinde tasarlanmış web arayüzüne aktarılmıştır. Aktarım sonucunda kullanıcıya daha güzel içerik sağlanmıştır. Oluşturulan sistem geliştirilmeye açıktır. İleride oluşabilecek yeni kavramların veri madenciliğinin niteliklerine eklenmesiyle daha kesin tespitler yapılabilir.

Uygulamanın arayüzü, kullanan banka için düzenlenerek kolay bir ekran sağlanmaktadır. Uygulamada oluşturulan arayüz ile bir müşteriye ait yeni bir alışveriş sisteme eklendiğinde arka planda Weka tarafından sağlanan karar ağacının çalışmasıyla dolandırıcılığın tespiti yapılmıştır.

Tüm yapı düşünüldüğü takdirde müşterinin beş aylık geçmişi ve alışveriş miktarları üzerinden alışveriş alışkanlığı belirlenerek yapılan yeni işlemlerde dolandırıcılık risk grubu ile karşılaşım sağlanmıştır.

15. KAYNAKLAR

Berry MJ., Linoff GS. (2004). Data Mining Techniques: For Marketing, Sales and Customer Relationship Management 2nd ed., Wiley, USA.

Bigus, JP. (1996) Data Mining With Neural Networks: Solving Business Problems from Application Development to Decision Support, McGraw Hill.

Bhardwaj A., Gupta R. (2016) Data miningbased detection – a comprehensive survey.

Özbay E. (2007) Finans Sektöründe Veri Madenciliği ile Dolandırıcılık Tespiti, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Konya, (Danışman: Prof.Dr. Ahmet Arslan).

Özkan Y. (2016) Veri Madenciliği Yöntemleri, Papatya Yayıncılık, 3. Basım, İstanbul.

Silahtaroglu G. (2008) Kavram ve Algoritmalarıyla Temel Veri Madenciliği, Papatya Yayıncılık, İstanbul.

Soylu K. (2018) Kredi Kartı Sahte İşlem Tespiti, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara, (Danışman: Prof.Dr. Şahin Emrah).

Şahin Y., Duman E. (2011) Detecting credit card fraud by decision trees and support vector machines.

Şeker ŞE. (2013) İş Zekası ve Veri Madenciliği Weka İle, 1. Basım, Cinius Yayıncılık, İstanbul.

Temel, GO., Çamdeviren, H., Akkuş, Z. (2005) Sınıflama Ağaçları Yardımıyla Restless Legs Syndrome (RLS) Hastalarına Tanı Koyma, İnönü Üniversitesi Tıp Fakültesi Dergisi.

<http://adudspace.adu.edu.tr:8080/xmlui/handle/11607/679> (01.02.2019)

<https://archive.ics.uci.edu/ml/datasets.php/> (09.04.2019)

<http://bilgisayarkavramlari.sadievrenseker.com/2009/06/01/weka/> (17.04.2019)

<http://bilgisayarkavramlari.sadievrenseker.com/2012/11/13/information-gain-bilgi-kazanimi/> (21.04.2019)

<https://ceaksan.com/tr/weka-nedir/> (20.04.2019)

<https://www.biltektasarim.com/blog/html-nedir> (11.03.2019)

<https://github.com/cagri34/Bankacilikta-Dolandiricilik-Yontemleri-ve-Tespiti/> (14.07.2019)

https://www.eecs.yorku.ca/course_archive/2006-07/W/4412/doc/weka (25.04.2019)

<https://www.mediaclick.com.tr/blog/mysql-nedir> (09.03.2019)

<https://www.mediaclick.com.tr/blog/php-nedir> (11.03.2019)

<https://www.pcnet.com.tr/makaleler/internet/internet-bankaciliginin-avantajlari/>
(16.01.2019)

<https://www.tarihnotlari.com/bankaciligin-tarihcesi/> (11.02.2019)

<https://www.tbb.org.tr/gec/KTPV14.pdf> (22.01.2019)

<https://veribilimcisi.com/2017/07/18/karisiklik-matrisi-nedir/> (14.05.2019)

<https://www.webprofu.com/css-faydalari-nelerdir> (01.05.2019)



16. ÖZGEÇMİŞ

ÇAĞRI CİVELEK

Aslen Trabzonlu olan Çağrı CİVELEK 1994 tarihinde İstanbul – Kadıköyde doğdu. 2000 – 2008 yılları arasında ilköğrenimini Hasan Leyli İlköğretim Okulu ‘nda tamamlamıştır. 2008 – 2012 yılları arasında ortaöğrenimini TEB Ataşehir Lisesi ‘nde tamamladı. 2012- 2016 yılları arasında Haliç Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği bölümünü tamamlamıştır. 2016 yılında Haliç Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı ‘nda Bilgisayar Mühendisliği Programına Tezli olarak başlamıştır.

