

## Karar Ağaçları ile Telefon Dolandırıcılığı Verilerinin Analizi

Erhan AKBAL<sup>1</sup>, Şengül DOĞAN<sup>2</sup>, Nurhayat VAROL<sup>3</sup>

<sup>1,2</sup>Fırat Üniversitesi Teknoloji Fakültesi Adli Bilişim Mühendisliği Bölümü, Elazığ, Türkiye

<sup>3</sup>Motorlu Araçlar ve Ulaştırma Teknolojileri Bölümü, Teknik Bilimler MYO, Fırat Üniversitesi, Elazığ, Türkiye  
<sup>1</sup>erhanakbal@firat.edu.tr, <sup>2</sup>sdogan@firat.edu.tr, <sup>3</sup>nvarol@firat.edu.tr

(Geliş/Received: 07.06.2016; Kabul/Accepted: 24.02.2017)

### Özet

İnsanlar teknolojinin gelişmesiyle beraber bankacılık, alışveriş, sosyal iletişim ve haberleşme gibi günlük pek çok işlerini sanal platformda bilgisayar, akıllı cihaz gibi araçlar yardımıyla gerçekleştirmektedir. Bu işlemler sanal ortamda gerçekleştirildikçe dolandırıcılık faaliyetleri de farklı boyutlar kazanmıştır. Özellikle küçük yaşlardan itibaren tüm insanların sürekli yanlarında bulundukları cep telefonları dolandırıcılar için yaygın uygulama aracı olmuştur. İnsanlarda bir anlık baskı, korku veya kişisel verilerin önceden izinsiz elde edilmesi sağlanarak resmi platformdan arama hissi yaratılması ile duygu karmaşası oluşturup insanların dolandırıcılık amaçlarına uygun hareket etmesi sağlanmaktadır. Bu çalışmada telefon dolandırıcılığı verilerinin analiz edilmesi sağlanmıştır. Kişilerin yaş, meslek grubu, yaşam yeri gibi bireysel özelliklerine göre telefon dolandırıcılığı verilerinin karar ağaçları yöntemlerinden C4.5 algoritması ile sınıflandırılması yapılmış ve sonuçlar değerlendirilmiştir.

**Anahtar Kelimeler:** Adli bilişim, Telefon dolandırıcılığı, Karar ağaçları, C4.5 algoritması, Veri analizi

## Analyzing Phone Fraud Data with Decision Trees

### Abstract

People realize many everyday tasks such as banking, shopping, social networking and communication with the development of technology by using tools as computer, smart devices etc. in virtual platform. The fraudulent activities have gained new dimensions with these processes in virtual environments. Mobile phones that people always carry with them have been common application tool for fraudsters. Fraudsters try to act in accordance with their fraudulent purposes providing a momentary pressure, fear or obtaining of personal data in an unauthorized manner. In this study, analysis of phone fraud data is performed. The phone fraud data are classified according to individual characteristics of people such as occupations, ages and living places by using C4.5 algorithm of decision tree method and the results are evaluated.

**Keywords:** Digital forensics, Phone fraud, Decision trees, C4.5 algorithm, Data analysis

### 1. Giriş

Teknolojinin büyük bir hızla gelişmesi insanların hızlı ve etkin bir şekilde günlük faaliyetlerini sanal platformda gerçekleştirme potansiyelini artırmaktadır [1]. İnsanlar özellikle internet tabanlı sistemleri kullanarak bankacılık işlemleri, rezervasyon yapma, evdeki elektronik cihazlara uzaktan müdahale etme gibi pek çok işlemi zaman ve mekân kavramı olmadan halledebilmektedir [2,3]. Bu işlemler için günümüzde hızlı bir şekilde kullanımı artan internet alt yapısına sahip bir mobil telefon, bilgisayar, tablet gibi cihazlar rahatlıkla kullanılabilir [4]. Özellikle mobil telefonlar insanların sürekli yanında taşıdıkları bir cihaz olduğundan ek bir yük gerektirmeden

teknolojiyi rahatlıkla takip etmemizi sağlamaktadır. İletişim kurma amacının dışında internet bankacılığı, e-gazete okuma, oyun oynama, sosyal platformları kullanma gibi pek çok amaçla yaygın kullanılan akıllı telefonlar dolandırıcılık faaliyetlerinde özellikle seçilen cihazların başında gelmektedir [5]. Böylelikle mevcut teknoloji insanların işlerini kolaylaştırmanın yanı sıra aynı zamanda yeni dolandırıcılık faaliyetleri ile yüz yüze gelmesine de neden olmuştur. Dolandırıcılar, telefon dolandırıcılığı faaliyetlerinde veri toplama, sosyal mühendislik ve resmi belgeleri ele geçirme yöntemlerini kullanarak kişi ile ilgili bilgileri elde ederek insanları amaçları doğrultusunda yönlendirmeye çalışır [5-7].

- Veri Toplama: İnsanların belli bir süre sanal ortamda takip edilmesi ile elde edilen verilerin depolanması sürecidir. Sosyal platformlarda yapılan mekân bildirimleri, yorumlar, beğeniler, ziyaret edilen alışveriş siteleri gibi başlıklar baz alınarak bir veri bloğu oluşturulup veri madenciliği teknikleri ile bireysel anlamlı bilgilere ulaşılmaya çalışılır.

- Sosyal Mühendislik: Sosyal mühendislik başlığı altında, veri toplama basamağında olduğu gibi kişilerle ilgili elde edilen bilgilerden belirgin özellikler kullanılıp insanların güveni kazanılarak zaaflarından yararlanılmaya çalışılır.

- Resmi Belgeleri Ele Geçirme: İnsanların resmi alanda kullandıkları belgeler ve dokümanların ele geçirilmesi suretiyle kişilerin resmi makamlarca aranyormuş izlenimi yaratılarak korku veya kafa karışıklığı yaratmak üzere amaca ulaşmaya çalışma sürecidir.

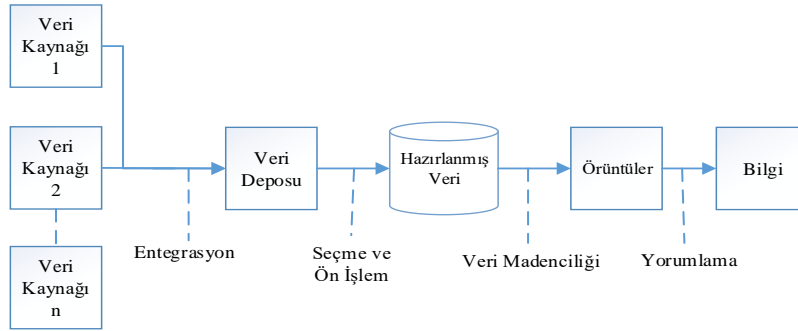
Yukarıda verilen maddelerin hepsinde temel amaç resmi veya bireysel verilerden elde edilen anlamlı bilgiler ile kişilerde güven, korku ya da kafa karışıklığı yaratarak dolandırıcılık faaliyetlerinde kişinin rıza göstermesi sürecinde zorlayıcı/özendirici etki oluşturma olduğu söylenebilir. Örneğin sosyal mühendislik/veri toplama süreçleri kapsamında kişinin x bankasında kredi kartı borcunu düzenli ödediği tespit edilirse “borcunuzu düzenli ödediğiniz için bankamızdan ..... hediyesini kazandınız” şeklinde mevcut banka adı, düzenli ödeme bilgileri ile kişinin güvenini kazanma yönünde dolandırıcılık faaliyetlerini gerçekleştirmeye yönelik mesajlar gönderilebilir. Bir diğer örnek

ise resmi makam isimleri ve unvanlar kullanılarak kişilerde korku ve tedirginlik yaratmak suretiyle dolandırıcılık faaliyetlerinin gerçekleştirilmesidir. Bu faaliyetlerin temelinde kişilerde ani korku, kafa karışıklığı veya kişilerle ilgili ulaşılması zor bilgilere ulaşarak güven ortamını yaratıp kişinin kendi rızası ile dolandırıcılık faaliyetinin başrolü olmasını sağlamaktır [6,7].

Bu çalışmada günümüzde sıkça karşılaşılan telefon dolandırıcılığı üzerine bir çalışmaya yer verilmiştir. Makale genel yapısında; veri madenciliği tekniklerine bölüm 2’de, karar ağaçları yapısına bölüm 3’de değinilmiştir. Bölüm 4 verilerin değerlendirildiği uygulamayı içermektedir. Bölüm 5’de ise sonuçlar değerlendirilmekte ve öneriler kısmı yer almaktadır.

## 2. Veri Madenciliği Teknikleri

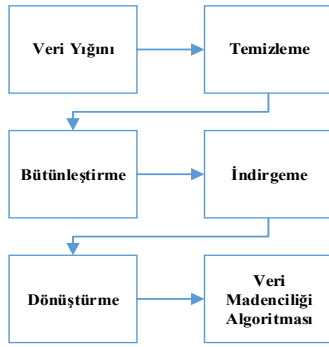
Günlük hayatta tıp, eğitim, endüstri, otomotiv, pazarlama gibi sektörler mevcut faaliyetlerini devam ettirenken aynı zamanda büyük veri yığınları oluşturmaktadır. Bu veri yığınları pek çok keşfedilmemiş bilgiyi bünyelerinde barındır [8]. Veri madenciliği, teknolojiye yaşanan hızlı değişikliklere paralel olarak her sektörde oluşan bu veri yığınları içerisinde farkındalık oluşturacak şekilde veriyi anlamlandırma süreci olarak tanımlanabilir. Veri madenciliğinde verinin bilgiye dönüştürülmesi süreci Şekil 1’de gösterilmiştir [9, 10].



Şekil 1. Veri Madenciliği Sürecinin Adımları

Şekil 1’de gösterildiği gibi bir veya daha fazla veri kaynağından elde edilen veri yığını belli standartlara uygun şekilde bir veri ambarında tutulur. Bu veri ambarındaki veriler Şekil 2’de gösterildiği gibi işlem basamaklarından

geçirilerek veri madenciliği tekniklerinin uygulanabileceği forma getirilir [9].



Şekil 2. Veri Ön İşlem Basamakları

Şekil 2 ile temsil edilen işlemlerden geçerek hazırlanan veri; veri madenciliği teknikleri ile işlenip yorumlanarak bilgi elde edilme süreci tamamlanır. Veri madenciliği; ilişkisel veri tabanları, istatistik/olasılık ve yapay öğrenme konularının tekniklerini kullanarak mevcut veriler ile verinin analiz edilmesi ve gerekirse gelecek durum tespiti yapılmasını sağlayabilmektedir [8].

Veri madenciliği sürecinde kullanılan modeller genellikle işlevlerine göre

- Sınıflandırma
- Kümeleme
- Birliktelik kuralları ve ardışık zamanlı örüntüler şeklinde gruplandırılabilir.

Sınıflandırma; veri tabanlarındaki verilerin değerlendirilmesi ile gizli örüntüleri ortaya çıkarmak amacıyla kullanılır. Sınıflandırma sürecinde Karar Ağaçları, Örnek Tabanlı Yöntemler, Yapay Sinir Ağları gibi pek çok metod kullanılmaktadır.

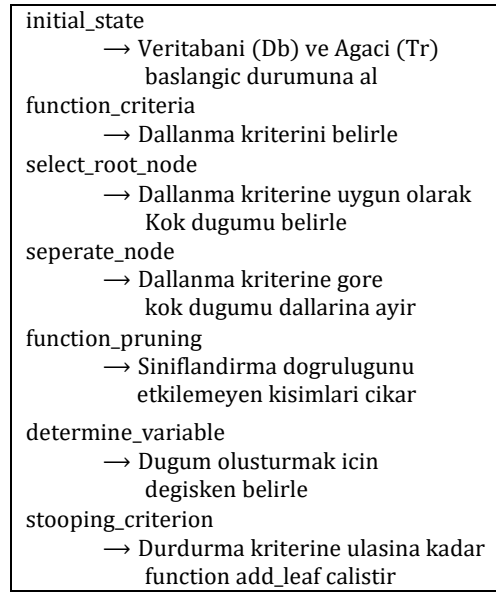
Kümeleme; verilerin paralel özelliklerinin tespit edilmesi ile gruplandırılması işlemidir. Kümeleme işleminde pek çok istatistiksel yöntem kullanılmaktadır. Kümeleme işleminde Bölme, Hiyerarşik Kümeleme yaygın kullanılan yöntemler arasındadır.

Birliktelik Kuralları ise veri tabanı içerisindeki yer alan mevcut verilerin birbirleriyle olan ilişkisinin tespit edilmesi olarak adlandırılmaktadır. Apriori Algoritması yaygın olarak kullanılmaktadır [8, 9].

### 3. Karar Ağaçları

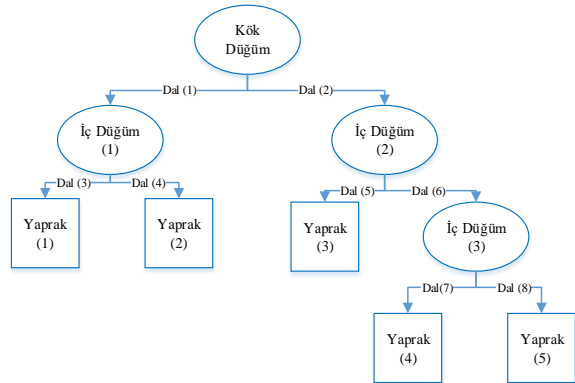
Sınıflandırma, veri yığınlarında verinin içerdiği ortak özelliklere göre gruplandırılması işlemidir. Sınıflandırma bir öğrenme alt yapısına dayanır. Öğrenmenin gerçekleşmesi için özel bir tahmin veya kural yapısına ihtiyaç duyulur.

Sınıflandırma işleminin verimli bir süreçte gerçekleştirilmesi için; veri yığınındaki verilerin yeterli sayıda, hatasız, sade ve tutarlı olması gerekir. Sınıflandırma işlemine yaygın kullanılan yöntemlerin başında karar ağacı yöntemi gelmektedir. Genel çalışma yapısı Şekil 3'te verilen karar ağaçları, veri yığınları içerisindeki verileri belirli kurallar çerçevesinde sınıflandırma ve küçük kümelere bölme işlemidir [10].



Şekil 3. Karar ağaçlarının genel çalışma yapısı

Karar ağacı Şekil 4'de gösterildiği gibi dal, düğüm ve yaprak olarak tanımlanan üç yapıdan oluşur. Karar ağacı yönteminde her bir dalın gerçekleşme olasılığı vardır. Böylelikle "kök düğümünden istenilen dala" veya "bulunulan daldan köke" ulaşılma olasılığı hesaplanabilir.



Şekil 4. Karar ağacı yapısı

Karar ağacı yönteminde veriler sınıflandırılırken veri yığını, eğitim ve test verisi olarak ikiye ayrılır ve eğitim verisi ile öğrenme işlemi sağlanır. Öğrenme kurallarının oluşturulması işlemi sonrasında test verisinin bu kurallar çerçevesinde sınıflandırılması sağlanır. Sonuçlar değerlendirilerek belirlenen kuralların başarımı test edilir. Başarı ölçütlerinin sağlanması durumunda kurallar yeni verilerin sınıflandırılması amacıyla kullanılabilir [11].

### 3.1. C4.5 algoritması

Karar ağacı modelinde geliştirilen ve yaygın kullanılan algoritmaların başında ID3 ve C4.5 algoritmaları gelmektedir. ID3 algoritması, eğitim veri kümesindeki kayıt ve veri tabanında nitelik sayısının fazla olması durumunda yoğun hesaplama yapmadan olası bir karar yapısı oluşturmak amacıyla kullanılır. C4.5 algoritması ise temelde ID3 algoritma mantığını kullanarak ID3 algoritmasındaki bazı eksik yönlerinin giderilmesi ile geliştirilmiştir [12]. Sadece kategorik veriler üzerinde çalışan ID3 algoritması Quinlan [13] tarafından sayısal değerler içeren niteliklerin kullanılmasını sağlayacak şekilde güncellenmiştir. C4.5 algoritmasında bilgi kazancı oranı öznelik seçim ölçütü olarak kullanılmaktadır. Böylelikle en yüksek bilgi kazancı oranına sahip öznelik seçilir [13,14]. C4.5 algoritmasının avantajları aşağıdaki gibi sıralanabilir.

- Hem kategorik hem de sayısal veri setleri kullanılabilir.
- Eğitim verilerinde eksik öznelik değerlerinin bulunması durumunda uygulanabilir.
- Eğitim veri setinin istenmeyen değerleri elimine edilebilir [15,16].
- Aşırı uygunluk problemini ortadan kaldırılır.

C4.5 algoritması, bir sistemin belirsizliği ve rastgeleliğin ölçüsü olarak tanımlanan entropi kavramına dayalı olarak işlem yapar. Eşitliği Denklem 1’de verilen entropide amaç veri kümesini sınıflara ayırmak için gerekli bilgi miktarını hesaplamaktır.

$$Entropi(S) = \sum_{i=1}^n -p_i \log_2(p_i) \quad (1)$$

Denklem 1’de,  $p_1, p_2, \dots, p_n$ ; veride hedef değişkenin sınıflarına ait olasılıklar, S ise veri

kümesi, n bir niteliğin sahip olabileceği farklı değerler miktarı olarak tanımlanmaktadır.

Karar ağaçları yöntemlerinde baskın (ayırt edici) niteliği belirlemek için her niteliğin bilgi kazancı hesaplanır. Bir A özelliğinin S örneği için bilgi kazancı Denklem 2’ye göre yapılmaktadır [12].

$$Kazanc(A, S) = Entropi(S) - \sum_{j=1}^n \frac{S_j}{S} Entropi(S_j) \quad (2)$$

Denklem 2 baz alınarak belirlenen her özellik için bilgi kazancı hesaplanır. Bilgi kazancı en yüksek olan özellik kök olarak tanımlanır. Bu işlemler her düğüm için

- Örneklerin aynı sınıfa ait olması
- Örneklerin bölünebilecek özelliklere ayıramaması
- Tüm özelliklerin uygun sınıfla temsil edilmesi gibi durumlardan biri oluşuncaya kadar devam eder [12,17].

S özneliği için bölünebilme bilgisi (B) denklem 3 ve kazanç oranı (K\_O) ise denklem 4 ile tanımlanmaktadır.

$$B = - \sum_{i=1}^n \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \quad (3)$$

$$K_O = \frac{Kazanc(A, S)}{B} \quad (4)$$

Karar ağaçlarında yapılan sınıflandırmada eğitim verisi tek bir sınıf içeren alt kümelere bölünebilir. Sonuçta karmaşık ve büyük bir ağaç yapısı elde edilir. Bu durumda budama olarak adlandırılan oluşturulacak ağaç yapısında alt kümelere ağaç yerine yaprak atanması ile ağaç yapısının gereksiz karmaşıklığı azaltılmış olur [18,19].

## 4. Yöntem

Bu makalede, 115 kişiden telefon dolandırıcılığı alanında elde edilen verilerin karar ağacı yöntemlerinden C4.5 algoritması ile değerlendirilmesi sağlanmıştır. 115 kişilik veri seti 75 eğitim ve 40 test verisi olarak kullanılmıştır. Kullanıcılara bireysel ve telefon dolandırıcılığı üzerine iki alanda sorular sorulmuştur. Bireysel sorular kategorize edilerek sistemdeki karşılıkları Tablo 1’de verilmiştir.

**Tablo 1.** Bireysel sorular ve sistemdeki eşdeğerleri

Bireysel Soru	Seçenekler	Kişi Dağılımları	Sistemdeki Karşılığı
Öğrenim Durumu	İlköğretim	8	OD_1
	Ortaöğretim	19	OD_0

(OD)	Lise	39	OD_L
	Üniversite	49	OD_U
	Köy	1	YY_K
Yaşam Yeri (YY)	İlçe	12	YY_I
	Şehir	94	YY_S
	Büyükşehir	8	YY_S
	Serbest Meslek	75	MES_S
Meslek (MES)	Memur	35	MES_M
	Kamu Güvenlik Görevlisi	5	MES_K
	20-30 arası	83	YS_1
	30-40 arası	21	YS_2
Yaş (YS)	40-50 arası	4	YS_3
	50-60 arası	6	YS_4
	60 yaş üstü	1	YS_5
	1000 TL altında	54	AYG_1
	1000-2000 TL arası	28	AYG_2
Aylık Gelir (AYG)	2000-3000 TL arası	30	AYG_3
	3000 TL üzeri	3	AYG_4

Sistemde telefon dolandırıcılığı üzerine belirlenen ve bireylere sorulan sorular ise aşağıda sunulmuştur.

- Kendinizin veya bir yakınınızın telefonla dolandırılması endişesi taşıyor musunuz?
- Telefon dolandırıcılığı sizin veya bir yakınınızın başına geldi mi?
- Telefon dolandırıcılığı sizin veya bir yakınınızın başına geldiyse gerekli yerlere şikâyette buldunuz mu bulduysanız sonuç alabildiniz mi?
- Sizi birisi arayarak Hâkim, Savcı veya Polis olarak tanıtp para istese ne yapardınız?
- Sizce Hâkim, Savcı ve Polis insanları arayarak para kontör talep eder mi?
- Telefon dolandırıcılığına karşı polisin bilgilendirme çalışmalarını yeterli buluyor musunuz?
- Telefon dolandırıcılığına yazılı ve görsel basında yeteri kadar yer veriliyor mu?
- Telefon dolandırıcılarına verilen cezaları yeterli buluyor musunuz?
- Sizce telefon dolandırıcılığının neden önüne geçilemiyor?
- Telefon dolandırıcılığının önüne geçmek için sizce ne yapılmalı?
- Telefon dolandırıcılığına karışanlara verilen ceza artırılırsa bu suçun önüne geçilebilir mi?
- Sizce insanlar neden telefonla dolandırıcılık yapmak isterler?
- Sizce insanlar neden bu dolandırıcılara inanır?

- Ülkemizde meydana gelen telefon dolandırıcılığını sayısal olarak nasıl değerlendiriyorsunuz?

Tablo 1’de verilen kısaltmalar baz alınarak bireysel sorular sisteme girdi olarak verilmiştir. Oluşturulan kural dizisinde insanların “Telefon dolandırıcılığı endişesi taşımaları” üzerine yapılan değerlendirmede sonuçlar Evet/Hayır şeklinde sınıflandırılmıştır. C4.5 algoritmasına göre her bir kural dizisinden elde edilen ortak kurallar Tablo 2’de verilmiştir.

Çalışmada oluşturulacak ağaç yapısı için daha az hesaplama içeren ve veri setinin sınıflandırılmasında en iyi yolu sorgulayan ön budama işlemi kullanılmıştır.

**Tablo 2.** Sistemden elde edilen ortak kural dizileri

Kural Dizisi	Sonuç
$if(OD=OD_I    OD=OD_O) \&\& (YY=YY_K    YY=YY_I    YY_S) \&\& (MES=MES_S    MES=MES_M    MES=MES_K) \&\& (YS=YS_I    YS=YS_2) \&\& (AYG=AYG_I    AYG=AYG_2)$	Hayır
$if(OD=OD_I    OD=OD_O) \&\& (YY=YY_K    YY=YY_I    YY_S) \&\& (MES=MES_S    MES=MES_M    MES=MES_K) \&\& (YS=YS_I    YS=YS_2) \&\& (AYG=AYG_3    AYG=AYG_4)$	Evet
$if(OD=OD_I    OD=OD_O) \&\& (YY=YY_K    YY=YY_I    YY_S) \&\& (MES=MES_S    MES=MES_M    MES=MES_K) \&\& (YS=YS_3    YS=YS_4    YS=YS_5) \&\& (AYG=AYG_I)$	Hayır
$if(OD=OD_I    OD=OD_O) \&\& (YY=YY_K    YY=YY_I    YY_S) \&\& (MES=MES_S    MES=MES_M    MES=MES_K) \&\& (YS=YS_3    YS=YS_4    YS=YS_5) \&\& (AYG=AYG_2    AYG=AYG_3    AYG=AYG_4)$	Evet
$if(OD=OD_L    OD=OD_U) \&\& (YY=YY_K    YY=YY_I    YY_S) \&\& (MES=MES_S    MES=MES_M    MES=MES_K) \&\& (YS=YS_I    YS=YS_2    YS=YS_3    YS=YS_4    YS=YS_5) \&\& (AYG=AYG_I    AYG=AYG_2    AYG=AYG_3)$	Hayır
$if(OD=OD_L    OD=OD_U) \&\& (YY=YY_K    YY=YY_I    YY_S) \&\& (MES=MES_S    MES=MES_M    MES=MES_K) \&\& (YS=YS_I    YS=YS_2    YS=YS_3    YS=YS_4    YS=YS_5) \&\& (AYG=AYG_4)$	Evet

Aynı zamanda çalışmada, kesinlik, duyarlılık, doğruluk ve F ölçümü kullanılarak sınıflama algoritmasının başarımı değerlendirilmiştir.

*Kesinlik (P):* Eşitliği denklem 5’te verilen sınıflandırma sonucunun kesinlik derecesi, pozitif olarak tanımlanan örneklerin pozitif olarak sınıflandırılan toplam örneklere oranı ile elde edilir.

$$P = \frac{TP}{TP + FP} \quad (5)$$

*Duyarlılık (R)*: Pozitif olarak tanımlanan örneklerin gerçekte toplam pozitif olanlara oranı duyarlılık olarak adlandırılır. Eşitliği Denklem 6'da verilmiştir.

$$R = \frac{TP}{TP + FN} \quad (6)$$

*Doğruluk (A)*: Denklem 7 ile tanımlanan doğruluk, doğru sınıflandırılan örneklerin toplam örnek sayısına oranı olarak ifade edilebilir.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

*F ölçümü (F)*: Sistemde kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır. Eşitliği denklem 8'de verilmiştir.

$$F = \frac{2 * R * P}{R + P} \quad (8)$$

Denklemlerde verilen eşitliklerde

*TP*; Pozitif olup sınıflandırma sonrasında da pozitif olarak sınıflandırılanlar

*TN*; Negatif olup sınıflandırma sonrasında da negatif olarak sınıflandırılanlar

*FP*; Pozitif olup sınıflandırma sonrasında pozitif olarak sınıflandırılmayanlar

*FN*; Negatif olup sınıflandırma sonrasında pozitif olarak sınıflandırılmayanlar şeklinde tanımlanmaktadır [16].

Belirlenen başarı ölçütlerinden elde edilen değerler Tablo 3'de verilmiştir.

**Tablo 3.** Sistemde elde edilen başarı ölçüt değerleri

	P	R	A	F
Eğitim	0.836	0.801	0.851	0.798
Test	0.813	0.797	0.840	0.791
Tümü	0.825	0.803	0.841	0.790
<b>Ortalama</b>	<b>0.824</b>	<b>0.800</b>	<b>0.844</b>	<b>0.793</b>

Ayrıca telefon dolandırıcılığı üzerine sistemde değerlendirilen 14 sorudan elde edilen tüm sınıflandırmalar incelendiğinde

- Dolandırıcılık faaliyetlerine maruz kalmayan insanların böyle bir durumda resmi prosedür göz önüne alındığında ne yapması gerektiğini bilmedikleri

- Dolandırıcılık faaliyetlerine maruz kalan insanların şikâyet sonrası bir çözüm bulamadıkları
- Telefon dolandırıcılığı kapsamında yapılan bilgilendirme çalışmalarının yeterli olmadığı
- İnsanların bu dolandırıcılık kapsamında ülkemizdeki cezaları bilmemelerine rağmen sık karşılaşılan bir durum olduğu için cezaların yeterli olmadığı kanaatinde oldukları
- Telefon dolandırıcılığı yapan kişileri tespit edecek yeterli yetişmiş insan gücünün mevcut olmamasından ötürü cezaların yeterli seviyeye çıkarılması durumunda bile bu suçların önlenmesinde sıkıntıların yaşanacağı sonuçlarına ulaşılmıştır.

## 5. Sonuçlar ve Öneriler

Bu çalışma, son günlerin popüler araştırma konularından biri olan telefon dolandırıcılığı üzerine bir uygulamayı içermektedir. Yapılan çalışmada, farklı bireysel özelliklere (öğrenim durumu, meslek grubu, yaş, yaşam yeri ve gelir düzeyi) sahip kişilerden telefon dolandırıcılığı üzerine elde edilen veriler karar ağaçları yöntemlerinden C4.5 algoritması ile sınıflandırılarak sonuçlar değerlendirilmiştir. Değerlendirme sonuçlarına göre telefon dolandırıcılığı insanların yaş, yaşam yeri, eğitim seviyesi vs. gibi bireysel özelliklerine bakılmaksızın yaşayabilecekleri ve her kişinin taşıdığı bir kaygı olarak gözlemlenmiştir.

Bu çalışmada elde edilen bulgular dikkate alındığında insanların dolandırıcılık faaliyetlerine maruz kalmamaları için yeterince bilgilendirilmeleri gerektiği ortaya çıkmaktadır. Aynı zamanda sanal ortamda yapılan dolandırıcılık faaliyetlerinin önlenmesinde devletin güvenlik ve gizlilik politikaları da büyük önem teşkil etmektedir. Kişisel bilgilerin korunması ve gizliliğinin muhafaza edilmesi yolunda devlet bünyesinde yetişecek insan gücünün ve yapılacak çalışmaların artırılması ile insanların gizli bilgilerinin üçüncü şahıslar tarafından ele geçmesinin önlenmesi sağlanabilmelidir. Bununla beraber telefon aracılığı ile yapılan dolandırıcılık faaliyetlerinde insanların bir anlık zafiyet, korku, endişe gibi duygu değişimlerinden faydalanılmaması için önceki yaşanan vakalarda tespit edilen başlıklar resmi platformlar üzerinden duyurulmalı ve

böylelikle insanların resmi platformlarda karşılaşabilecekleri ve karşılaşmayacakları durumlar üzerine bir bilinç oluşturulması sağlanabilmelidir.

## 6. Kaynaklar

1. Wales, E. (2003). E-commerce counts cost of online card fraud. *Computer Fraud & Security*, **1**, 9-11.
2. Akçetin, E., Çelik, U., Üniversitesi, B., & Programı, B. P. (2015). A classification analysis on the direct marketing campaigns in telephone banking with ant colony optimization classification algorithm method, *Journal of Internet Applications and Management*, **6(1)**, 5-19.
3. Blanco Hache, A. C., & Ryder, N. (2011). 'Tis the season to (be jolly?) wise-up to online fraudsters. Criminals on the Web lurking to scam shoppers this Christmas: 1 a critical analysis of the United Kingdom's legislative provisions and policies to tackle online fraud. *Information & Communications Technology Law*, **20(1)**, 35-56.
4. Chae, M., Shim, S., Cho, H., & Lee, B. (2007, January). An empirical analysis of fraud detection in online auctions: Credit card phantom transaction. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, 155a-155a, IEEE.
5. Thomas, B., Clergue, J., Schaad, A., & Dacier, M. (2004). A comparison of conventional and online fraud. In *CRIS*, **4**, 25-27.
6. Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical science*, 235-249.
7. Brause, R., Langsdorf, T., & Hepp, M. (1999). Neural data mining for credit card fraud detection. In *Tools with Artificial Intelligence, 1999. Proceedings. 11th IEEE International Conference on*, pp. 103-106, IEEE.
8. Rokach, L., & Maimon, O. (2014). *Data mining with decision trees: theory and applications*. World Scientific.
9. Tan, P. N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*, **1**, Boston: Pearson Addison Wesley.
10. Ozkan, Y. (2008). *Veri Madenciliği Yöntemleri*, Papatya Yayıncılık.
11. Sarmanova, A., & Albayrak, S. (2013, April). Alleviating class imbalance problem in data mining. In *Signal Processing and Communications Applications Conference (SIU), 2013 21st*, 1-4, IEEE.
12. Bahety, A. (2014). Extension and Evaluation of ID3–Decision Tree Algorithm. *Entropy (S)*, **2**, 1.
13. Şatır, E., Azboy, F., Aydın, A., Arslan, H., & Hacıfendioğlu, Ş. (2016). Veri İndirgeme ve Sınıflandırma Teknikleri ile Glokom Hastalığı Teşhisi. *El-Cezeri Journal of Science and Engineering*, **3(3)**, 485-497.
14. Bulut, F. (2016). Çok Katmanlı Algılayıcılar ile Doğru Meslek Tercihi. *Anadolu University Journal Of Science And Technology–A Applied Sciences and Engineering*, **17(1)**, 97-109.
15. Ture, M., Tokatli, F., & Kurt, I. (2009). Using Kaplan–Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4. 5 and ID3) in determining recurrence-free survival of breast cancer patients. *Expert Systems with Applications*, **36(2)**, 2017-2026.
16. Kaynar, O., Görmez, Y., Yıldız, M., & Albayrak, A. (2016). Makine Öğrenmesi Yöntemleri ile Duygu Analizi. *International Artificial Intelligence and Data Processing Symposium (IDAP'16)*, September 17-18, Malatya/TURKEY, 234-241.
17. Adhatrao, K., Gaykar, A., Dhawan, A., Jha, R., & Honrao, V. (2013). Predicting Students' Performance using ID3 and C4. 5 Classification Algorithms. *arXiv preprint arXiv:1310.2071*.
18. Cheng, J., Fayyad, U. M., Irani, K. B., & Qian, Z. (1988). Improved decision trees: a generalized version of id3. In *Proc. Fifth Int. Conf. Machine Learning*, 100-107.
19. Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote sensing of environment*, **61(3)**, 399-409.