



MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



**TUŞA BASIŞ DİNAMİKLERİ,
AKSELEROMETRE VE JİROSKOP
VERİLERİYLE MOBİL CİHAZLARDA
DAVRANIŞSAL BİYOMETRİK
KİMLİK DOĞRULAMA**

UMUT BERHAN BALKIR

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği

Anabilim Dalı

Bilgisayar Mühendisliği Programı

DANIŞMAN

Dr. Zehra Aysun ALTIKARDEŞ

İSTANBUL, 2020



MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



**TUŞA BASIŞ DİNAMİKLERİ,
AKSELEROMETRE VE JİROSKOP
VERİLERİYLE MOBİL CİHAZLARDA
DAVRANIŞSAL BİYOMETRİK
KİMLİK DOĞRULAMA**

UMUT BERHAN BALKIR
(523616022)

YÜKSEK LİSANS TEZİ
Bilgisayar Mühendisliği
Anabilim Dalı
Bilgisayar Mühendisliği Programı

DANIŞMAN
Dr. Zehra Aysun ALTIKARDEŞ

İSTANBUL, 2020

MARMARA ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

Marmara Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Öğrencisi Umut Berhan BALKIR'ın "Tuşa Basış Dinamikleri, Akselerometre ve Jiroskop Verileriyle Mobil Cihazlarda Davranışsal Biyometrik Kimlik Doğrulama" başlıklı tez çalışması, 22 Ocak 2020 tarihinde savunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Dr. Öğr. Üyesi Zehra Aysun ALTIKARDEŞ (Danışman)

Marmara Üniversitesi Teknik Bilimler Meslek Yüksek Okulu Bilgisayar Teknolojileri Bölümü

(İMZA).....

Prof. Dr. Ali BULDU (Üye)

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü

(İMZA).....

Doç. Dr. Muhammed Ali AYDIN (Üye)

İstanbul Üniversitesi Cerrahpaşa Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü

(İMZA).....

ONAY

Marmara Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 05.02.2020 tarih ve 2020/05-03 sayılı kararı ile Umut Berhan BALKIR'ın Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Programında Yüksek Lisans derecesi alması onanmıştır.

Fen Bilimleri Enstitüsü Müdürü

Prof. Dr. Bülent EKİCİ




TEŞEKKÜR

Öğrenim hayatımda ve tez çalışmamda bilgi ve tecrübelerinden yararlandığım değerli hocam ve tez danışmanım Sayın Dr. Öğr. Üyesi Zehra Aysun ALTIKARDEŞ'e, manevi desteklerinden ötürü eşime, anneme, babama ve kardeşime teşekkürlerimi sunarım.

Ocak 2020

Umut Berhan BALKIR



İÇİNDEKİLER

1. GİRİŞ.....	1
1.1. Biyometrik Kimlik Doğrulama.....	1
1.2. Davranışsal Biyometrik Kimlik Doğrulama.....	2
1.3. Literatür Araştırmaları.....	3
1.3.1. Davranışsal Biyometrik Kimlik Doğrulama Alanında Yapılan Araştırma Çalışmaları.....	3
1.3.2. Klasik Biyometrik Kimlik Doğrulama Yöntemleri İçin Yapılan Siber Tehdit Çalışmaları.....	3
1.3.3. Bilgisayar Ortamında Tuşa Basış Dinamikleriyle Yapılan Kimlik Doğrulama Çalışmaları.....	4
1.3.4. Mobil Cihazlarda Tuşa Basış Dinamikleriyle Yapılan Kimlik Doğrulama Çalışmaları.....	5
1.3.5. Mobil Cihazlarda Sensör Verileriyle Yapılan Kimlik Doğrulama Çalışmaları.....	6
1.3.6. Mobil Cihazlarda Tuşa Basış Dinamikleri ve Sensör Verileriyle Yapılan Kimlik Doğrulama Çalışmaları.....	9
1.3.7. Mobil Cihazlardaki Sensörler İçin Yapılan Siber Tehdit Çalışmaları.....	10
2. MATERYAL VE YÖNTEM.....	11
2.1. Tuşa Basış Dinamikleri ve Sensör Verileri.....	11
2.1.1. Zaman.....	11
2.1.2. Akselerometre.....	12
2.1.3. Jiroskop.....	12
2.2. Veri ve Toplanması.....	12
2.2.1. Veri Toplama Uygulaması.....	12
2.2.2. Veri Setinin Oluşturulması.....	14

2.2.3.	Veri Setinin Demografik Yapısı	16
2.3.	Verilerin Analizi	16
2.4.	Analiz Yöntemi	17
2.4.1.	Makine Öğrenmesi	17
2.4.1.1.	Karar Ağacı Sınıflandırıcı (Decision Tree Classifier).....	18
2.4.1.2.	Rastgele Orman Sınıflandırıcı (Random Forest Classifier)	18
2.4.1.3.	Naive Bayes Sınıflandırıcı (Naive Bayes Classifier)	18
2.4.1.4.	Yapay Sinir Ağları (Neural Network).....	18
2.4.1.5.	Yapay Sinir Ağları - Temel Bileşen Analizi (Neural Network – Principle Component Analysis).....	18
2.4.2.	Performans Göstergeleri	19
2.4.2.1.	Yanlış Reddetme Oranı (False Rejection Rate) (FRR)	19
2.4.2.2.	Yanlış Kabul Oranı (False Acceptance Rate) (FAR)	20
2.4.2.3.	Alıcı İşletim Karakteristiği (Receiver Operating Characteristic) (ROC) Eğrisi ve Eğri Altındaki Alan (Area Under Curve)(AUC).....	20
2.4.2.4.	Eş Hata Oranı (Equal Error Rate) (EER)	22
2.4.2.5.	Doğruluk Oranı (Accuracy)	23
3.	BULGULAR VE TARTIŞMA.....	24
3.1.	Tüm Sonuçlar.....	24
3.2.	Ortalama Analiz Sonuçları.....	28
3.2.1.	Karar Ağacı Sınıflandırıcı Çalışmaları	28
3.2.1.1.	Zaman Veri Seti	28
3.2.1.2.	Zaman ve Akselerometre Veri Seti	29
3.2.1.3.	Zaman ve Jiroskop Veri Seti	30
3.2.1.4.	Zaman, Akselerometre ve Jiroskop Veri Seti.....	31

3.2.2.	Rastgele Orman Sınıflandırıcı Çalışmaları.....	32
3.2.2.1.	Zaman Veri Seti	32
3.2.2.2.	Zaman ve Akselerometre Veri Seti	33
3.2.2.3.	Zaman ve Jiroskop Veri Seti	34
3.2.2.4.	Zaman, Akselerometre ve Jiroskop Veri Seti.....	35
3.2.3.	Naive Bayes Sınıflandırıcı Çalışmaları	36
3.2.3.1.	Zaman Veri Seti	36
3.2.3.2.	Zaman ve Akselerometre Veri Seti	37
3.2.3.3.	Zaman ve Jiroskop Veri Seti	38
3.2.3.4.	Zaman, Akselerometre ve Jiroskop Veri Seti.....	39
3.2.4.	Yapay Sinir Ağları Çalışmaları	40
3.2.4.1.	Zaman Veri Seti	40
3.2.4.2.	Zaman ve Akselerometre Veri Seti	41
3.2.4.3.	Zaman ve Jiroskop Veri Seti	42
3.2.4.4.	Zaman, Akselerometre ve Jiroskop Veri Seti.....	43
3.2.5.	Yapay Sinir Ağları - Temel Bileşen Analizi Çalışmaları	44
3.2.5.1.	Zaman Veri Seti	46
3.2.5.2.	Zaman ve Akselerometre Veri Seti	47
3.2.5.3.	Zaman ve Jiroskop Veri Seti	48
3.2.5.4.	Zaman, Akselerometre ve Jiroskop Veri Seti.....	49
3.3.	Ortalama Sonuçların Özeti.....	50
3.4.	Örneklem Sayına Göre Sonuç Karşılaştırma Çalışması	50
3.5.	Çoklu Sınıflandırma Deneysel Çalışması	52
4.	SONUÇLAR.....	53
4.1.	Sonuçların Değerlendirilmesi	53

4.1.1.	Farklı Veri Seti Kombinasyonlarının Sonuca Etkisi	53
4.1.2.	Farklı Makine Öğrenmesi Algoritmalarının Sonuca Etkisi	54
4.1.3.	En İyi Sonuçların Değerlendirilmesi	55
4.1.4.	Örneklem Sayısına Göre Sonuçların Değerlendirilmesi	55
4.1.5.	Çoklu Sınıflandırma Deneysel Çalışmasının Değerlendirilmesi.....	55
4.2.	Siber Tehditler	56
4.3.	Kullanılabilecek Alanlar	57
4.4.	Gelecekteki Çalışmalar	57



ÖZET

TUŞA BASIŞ DİNAMİKLERİ VE AKSELEROMETRE VERİLERİYLE MOBİL CİHAZLARDA KİMLİK DOĞRULAMA

Günümüzde sıkça kullandığımız ve birçok ihtiyacın karşılanabildiği mobil cihazlarda, bilinen anahtarla yani şifreyle kimlik doğrulama yöntemine ek olarak veya doğrudan bu yöntemin yerine, parmak izi ve yüz tanıma sistemleri gibi biyometrik kimlik doğrulama yöntemleri yaygın bir şekilde kullanılmaktadır. Bu kimlik doğrulama yöntemleri güvenilir görünseler de doğrulamayı sağlayan biyometrik veriler gizlenememektedir. Parmak izi, ses ve yüz gibi fiziksel biyometrik veriler sürekli olarak bir yerde bırakılmaktadır veya başkalarından saklanamamaktadır. Başka bir deyişle herkes tarafından görülen, duyulan, ulaşılabilen ve değiştirilemeyen verilerle kimlik doğrulama işlemleri gerçekleştirilmektedir. İleride bu verilerin çalınarak ya da kopyalanarak, doğrulama sistemlerinin kolayca atlatılamayacağı meçhul. Bu yöntemlerin yerine veya tercihe bağlı olarak bu yöntemlere ek olacak şekilde kullanılabilecek davranışsal biyometrik kimlik doğrulama yöntemleriyle, kullanıcıyı ek bir doğrulama adımıyla uğraştırmadan, arka planda yapılan davranış analizleri sayesinde belirli başarı yüzdeleriyle kimlik doğrulama yapılabilir. Bu çalışmada, çalınması ve taklit edilmesi çok daha zor gözükse, arka planda transparan bir şekilde çalıştığı için de kullanıcı deneyimi açısından daha verimli olan ‘tuşa basış dinamikleri ve mobil cihazlarda bulunan akselerometre ve jiroskop sensörleriyle elde edilen üç boyutlu uzaydaki tuşlama esnasındaki cihaz pozisyonu verileriyle’ davranışsal biyometrik kimlik doğrulama çözümü ele alınmıştır. Analizler birden fazla makine öğrenmesi algoritması ve farklı veri seti kombinasyonları kullanılarak yapılmıştır. Ortalama sonuçlara göre en iyi %1.30 eş hata oranı (EER) Rastgele Orman sınıflandırıcı elde edilmiştir.

ABSTRACT

KEYSTROKE DYNAMICS AND ACCELEROMETER DATA BASED USER AUTHENTICATION ON MOBILE DEVICES

In mobile devices which we use frequently and can handle many high-importance works, fingerprint or face recognition systems are widely used as biometric authentication methods in directly use or as additional factors with the password authentication methods. Although these authentication methods appear to be reliable, biometric data that provides authentication cannot be hidden. Physical biometric data such as fingerprints, sounds and facials are permanently left in places or cannot be hidden from others. In other words, authentication is performed with data that is seen, heard, accessible by everyone else and these are unchangeable. It is not known whether verification systems can be easily circumvented by stealing or copying this data in the future. Behavioral biometric authentication methods can be used in place of or optionally in addition to these methods, can be performed with certain success percentages by means of background behavior analyzes without the user having to deal with an additional verification step. In this study discussed a biometric authentication solution named 'keystroke dynamics and device position data during keying in three-dimensional space obtained by accelerometer and gyroscope sensors on mobile devices' which is much more difficult to steal and imitate and which is more efficient in terms of user experience because it works transparently in the background. The analyzes were performed using different machine learning algorithms and different data set combinations. According to the average results, the best equal error rate of 1.30% was obtained with Random Forest Classifier.

SEMBOLLER

m : Metre, uzunluk ölçü birimi

s : Saniye, zaman ölçü birimi

rad : Radyan, açı ölçü birimi

s² : Saniyenin karesi

% : Yüzdelik



KISALTMALAR

YSA	: Yapay Sinir Ağları
EER	: Equal Error Rate
FAR	: False Acceptance Rate
FRR	: False Reject Rate
ROC	: Receiver Operating Characteristic
AUC	: Area Under Curve
PCA	: Principal Component Analysis
TPR	: True Positive Rate
FPR	: False Positive Rate
ORT	: Ortalama
AKS	: Akselerometre
JIR	: Jiroskop
JSON	: Javascript Object Notation

ŞEKİL LİSTESİ

Şekil 1 Mobil Uygulama Veri Giriş Sayfası.....	13
Şekil 2 Mobil Uygulama Veri Görüntüleme ve Gönderme Sayfası.....	13
Şekil 3 Analiz Yöntemi Akış Diagramı.....	17
Şekil 4 Örnek ROC Eğrisi	21
Şekil 5 Örnek ROC Eğrisi Üzerinde EER Noktası.....	22
Şekil 6 Karar Ağaçları Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası	28
Şekil 7 Karar Ağaçları Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası.....	29
Şekil 8 Karar Ağaçları Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	30
Şekil 9 Karar Ağaçları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası.....	31
Şekil 10 Rastgele Ormanlar Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası.....	32
Şekil 11 Rastgele Ormanlar Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası.....	33
Şekil 12 Rastgele Ormanlar Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	34
Şekil 13 Rastgele Ormanlar Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	35
Şekil 14 Naive Bayes Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası	36
Şekil 15 Naive Bayes Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası	37

Şekil 16 Naive Bayes Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	38
Şekil 17 Naive Bayes Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası.....	39
Şekil 18 Yapay Sinir Ağları Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası.....	40
Şekil 19 Yapay Sinir Ağları Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası.....	41
Şekil 20 Yapay Sinir Ağları Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	42
Şekil 21 Yapay Sinir Ağları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	43
Şekil 22 Zaman Veri Seti Temel Bileşen Sayısı ve Kümülatif Toplam Varyans Grafiği	44
Şekil 23 Zaman ve Akselerometre Veri Seti Temel Bileşen Sayısı ve Kümülatif Toplam Varyans Grafiği	44
Şekil 24 Zaman ve Jiroskop Veri Seti Temel Bileşen Sayısı ve Kümülatif Varyans Grafiği.....	45
Şekil 25 Zaman, Akselerometre ve Jiroskop Veri Seti Temel Bileşen Sayısı ve Kümülatif Varyans Grafiği.....	45
Şekil 26 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası.....	46
Şekil 27 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası.....	47
Şekil 28 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	48
Şekil 29 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası	49

TABLO LİSTESİ

Tablo 1 Girilen Metin Harf Kısaltmaları.....	14
Tablo 2 Toplanan Verilerin Açıklamaları ve Kısaltmaları.....	14
Tablo 3 Veri Seti Parametreleri ve Değer Tipleri.....	15
Tablo 4 Veri Seti Dağılımı	16
Tablo 5 FAR Değerinin Tespiti.....	19
Tablo 6 FRR Değerinin Tespiti	20
Tablo 7 ROC Eğrisinin Tespiti.....	21
Tablo 8 Tüm Analiz Sonuçları	24
Tablo 9 Karar Ağaçları Algoritmasıyla Zaman Veri Setinin Sonuçları.....	28
Tablo 10 Karar Ağaçları Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları.....	29
Tablo 11 Karar Ağaçları Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları.....	30
Tablo 12 Karar Ağaçları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları.....	31
Tablo 13 Rastgele Ormanlar Algoritmasıyla Zaman Veri Setinin Sonuçları.....	32
Tablo 14 Rastgele Ormanlar Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları.....	33
Tablo 15 Rastgele Ormanlar Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları	34
Tablo 16 Rastgele Ormanlar Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları	35
Tablo 17 Naive Bayes Algoritmasıyla Zaman Veri Setinin Sonuçları.....	36
Tablo 18 Naive Bayes Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları	37

Tablo 19 Naive Bayes Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları	38
Tablo 20 Naive Bayes Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları.....	39
Tablo 21 Yapay Sinir Ağları Algoritmasıyla Zaman Veri Setinin Sonuçları.....	40
Tablo 22 Yapay Sinir Ağları Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları.....	41
Tablo 23 Yapay Sinir Ağları Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları	42
Tablo 24 Yapay Sinir Ağları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları	43
Tablo 25 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman Veri Setinin Sonuçları	46
Tablo 26 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları.....	47
Tablo 27 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları	48
Tablo 28 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları.....	49
Tablo 29 Ortalama Sonuçların Özeti	50
Tablo 30 Örneklem Sayısına Göre Sonuçlar	51
Tablo 31 Çoklu Sınıflandırma Deneysel Çalışmasının Sonuçları	52
Tablo 32 Özniteliklerin Sonuca Olan Etkileri	53
Tablo 33 Sınıflandırma Algoritmalarının Performansları	54

1. GİRİŞ

Günümüzde mobil cihazlar birçok kişinin hayatında yüksek bir öneme sahiptir. Kullanıcılar uzun zamandır haberleşme, sosyal medya iletişimi ve eğlence uygulamaları gibi konularda günlük işlerinin ve ihtiyaçlarının birçoğunu bu cihazlar vasıtasıyla hızlıca halledebilmektedirler. Buna paralel olarak bankacılık ve finans uygulamaları, tıbbi kimlik uygulamaları, kişisel bulut veri depolama çözümleri, e-ticaret, kayıtlı ödeme altyapıları gibi kritik ve hassas kimlik doğrulama gerektiren uygulamalar da mobil çözümler arasında yaygınlaştı. Böylece mobil cihazlarda güvenlik ihtiyacı günden güne daha da önem kazanmaya başladı. Ortamın mobil olması ve kolay ulaşılabilirliği göz önünde bulundurulduğunda, güvenlik gereksinimlerinin en başında kimlik doğrulama ihtiyacı ortaya çıkmış ve güvenlik alanında bir takım gelişmeler yaşanmıştır. Önceleri yalnızca bilinen anahtarlar üzerinden (parola, tek kullanımlık şifre) yürütülen kimlik doğrulama yöntemlerinin yerini, daha güvenli ve yetkisiz kişilerce elde edilmesi daha zor olan biyometrik kimlik doğrulama yöntemleri almıştır.

1.1. Biyometrik Kimlik Doğrulama

Biyometrik kimlik doğrulama, bir sisteme erişmeye çalışan kişinin, erişime yetkili olup olmadığını doğrulamak için kullanıcının biyometrik verilerinin toplanmasını ve doğrulanmasını sağlayan güvenlik yöntemidir. Biyometrik veriler, kişiye özgü olan kolay karşılaştırılabilen fiziksel ve biyolojik özelliklerdir. Sisteme erişmeye çalışan bir kullanıcının biyometrik verileri, yetkilendirmiş kişilerin biyometrik verileriyle örtüşüyorsa, sisteme giriş izni verilir.

Biyometrik kimlik doğrulama sistemleri bilgisayarlarda, akıllı telefonlarda ve tabletlerde kullanıldığı gibi devlet kurumlarında, özel şirketlerde ve sınır kapılarında da yaygın olarak kullanılmaktadır. Bilinen ve kullanılan bazı biyometrik kimlik doğrulama yöntemlerine aşağıdaki örnekler verilebilir [1];

- Yüz tanıma
- Parmak izi doğrulama
- Retina tarama
- İris tarama
- Ses Analizi

Günümüz mobil cihazlarında Parmak İzi Doğrulama ve Yüz Tanıma yöntemleri yaygın olarak kullanılmaktadır.

1.2. Davranışsal Biyometrik Kimlik Doğrulama

Biyometrik kimlik doğrulamanın bir alt konusu olan davranışsal biyometrik kimlik doğrulama, insanların gerçekleştirmiş oldukları faaliyetlerine ya da alışkanlıklarına bağlı olarak toplanan örüntülerin doğrulanması yöntemiyle çalışan bir kimlik doğrulama yöntemidir.

Davranışsal olmayan biyometrik kimlik doğrulama yöntemlerinin en büyük ortak dezavantajı, doğrulama sırasında kullanılan biyometrik verilerin gizlenememesi ve genele açık olmasıdır. Bu veriler için kişilerin iradesi dışında kopyalanması ve kullanılmasına ilişkin riskler mevcuttur. Ek olarak, literatürde bu yöntemlerin atlatılmasıyla ilgili çalışmalar mevcuttur [1-3]. Bu çalışmalarla ilgili detaylara (bkz: Bölüm 1.3) yer verilmiştir. Bunlarla beraber, kişilerin her doğrulama adımında; bu verileri kusursuzca sunabilmesi gerekmektedir. Örneğin parmak izi ile doğrulama yapan sistemlerde, parmağın ıslanması veya kirlenmesi gibi durumlarda birden fazla kez parmak izi doğrulama denemesi gerekebilmektedir. Bu durum da kullanıcı deneyimi açısından zaman zaman zorluklar doğurabilmektedir.

Günümüz kimlik doğrulama yöntemlerine alternatif olarak, taklit edilmesi daha zor olan, insanların davranışlarına ve alışkanlıklarına göre yapılan davranışsal biyometrik kimlik doğrulamalar, klasik biyometrik kimlik doğrulama yöntemlerine kıyasla öne çıkmaktadır. Aynı zamanda kimlik doğrulama için kullanıcıyla doğrudan bir etkileşim kurulmadığı ve arka planda transparan bir şekilde çalışıldığı için kullanıcı deneyimi de iyileştirilmektedir.

Tüm bunlara ek olarak birçok işlemin aynı seviyede kimlik doğrulama seviyesi gerektirmediği yaklaşımıyla hareket edilerek, gerçekleştirilecek işlemin risk seviyesine uygun oranda biyometrik kimlik doğrulama eşiği belirlenebilmektedir. Örneğin bir bankacılık uygulamasında gerçekleştirilecek para transferi limitleri, davranışsal biyometrik kimlik doğrulama sonuçlarına paralel olarak belirlenebilir. Bu tarz kullanımların getirdiği esneklik ile davranışsal biyometrik kimlik doğrulama

yöntemlerinin hem kullanıcı hem de uygulama sağlayıcısı için daha avantajlı ve kullanışlı bir yöntem olduğu görülmektedir.

Bu tez çalışmasında, insanların davranışlarına ve alışkanlıklarına göre türetilen biyometrik kimlik doğrulama yöntemi olarak; **‘tuşa basış dinamikleri, mobil cihazlarda bulunan akselerometre ve jiroskop sensörleriyle elde edilen üç boyutlu uzaydaki tuşlama pozisyonu verileriyle’** davranışsal biyometrik kimlik doğrulama çözümü Makine Öğrenmesi yöntemleri kullanılarak ortaya konmuştur.

1.3. Literatür Araştırmaları

Literatürdeki çalışmalar, bu tezin konusunun daha iyi yorumlanabilmesi için, aşağıdaki başlıklar kapsamında incelenmiştir.

1.3.1. Davranışsal Biyometrik Kimlik Doğrulama Alanında Yapılan Araştırma Çalışmaları

Mahfouz A. ve arkadaşları makalelerinde, mobil cihazlarda davranışsal biyometrik kimlik doğrulama konusunda geniş bir araştırmaya yer vermişlerdir. Genel olarak bu sistemlerin nasıl tasarlandığından bahsettikleri görülmektedir. Kullanılabilecek birçok özelliği belirtip, akselerometre ve jiroskop sensörlerine de değinmişlerdir [4].

Gümüş F. ve arkadaşları makalelerinde, davranışsal biyometrik kimlik doğrulama kavramı, ne gibi özelliklerle davranışsal biyometrik doğrulama yapılabileceği, bu doğrulama yönteminin avantajları ve dezavantajlarından bahseden bir inceleme çalışmasına yer vermişlerdir. Mobil cihazlarda bulunan sensörler vasıtasıyla davranışsal biyometrik kimlik doğrulama yapılabildiğine değinmişlerdir [1].

Alotaibi S. ve arkadaşları konferans bildirilerinde, literatürde yer alan davranışsal biyometrik kimlik doğrulama yöntemlerinin araştırmasına yer vermişlerdir. En etkili yöntemin, birden fazla özelliğin birleştirilerek yapıldığı analizler olduğunu belirtmişlerdir [5].

1.3.2. Klasik Biyometrik Kimlik Doğrulama Yöntemleri İçin Yapılan Siber Tehdit Çalışmaları

Galbally J. ve arkadaşları makalelerinde, yüz tanıma sistemlerini atlatmaya yönelik analizler yapmışlardır. Gerçek kişilerin yüz şekillerini vektörler halinde sahte kişilerin

yüzlerine aktararak %85'in üzerinde başarı oranıyla sistemleri atlatabilmişlerdir [2]. Böylece, bu durum her yüz tanıma sistemi için geçerli olmasa da yüz tanıma sistemlerinin atlatılabileceği ispatlanmıştır.

Espinoza M. ve arkadaşları makalelerinde, farklı özelliklerde plastik materyaller kullanılarak gerçek kişilerin parmak izlerinin toplanıp bu materyallere aktarılmasının sonrasında; bu materyallerin sahte kişilerin parmaklarına yerleştirilmesi ile parmak izi tanıma sistemlerinin verdiği başarı skorlarını incelemişlerdir. Gerçek parmak izlerinin verdiği skorlar bütün analizlerde daha yüksek olsa da, doğrulama oranının gerçek parmaklardan toplanıp plastik materyallere aktarılan ve sonrasında sahte kişilerin parmaklarına yerleştirilen parmak izlerinin skorlarıyla ayırt edilebilecek düzeyde olmadığı kanısına varılmıştır. Hatta bazı kişilerden kopyalan parmak izlerinin, orijinal parmak izlerinden daha yüksek başarı skoru verdiği gözlemlenmiştir [3].

1.3.3. Bilgisayar Ortamında Tuşa Basış Dinamikleriyle Yapılan Kimlik Doğrulama Çalışmaları

Özen Z. yaptığı doktora tezi çalışmasında, bilgisayar ortamında yazılan bir Java uygulaması vasıtasıyla kullanıcılardan güçlü şifreyi temsil eden metinler için tuşa basış dinamikleri verilerini toplamış, Yapay Sinir Ağları (YSA) yöntemiyle kimlik doğrulama çalışması yapmıştır. Çeşitli YSA algoritmalarıyla yapılan analizler sonucunda en iyi %2.12 Eş Hata Oranı (EER) oranı tespit edilmiştir [6].

Agun N. yaptığı yüksek lisans tezi çalışmasında, bilgisayar ortamında geliştirilen uygulama vasıtasıyla alan tabanlı tuşa basış dinamikleri çalışması yapmıştır. Klavyeler üzerinde bazı tuşlar birden fazla olmak üzere farklı konumlarda bulunabilmektedirler (Örneğin kontrol tuşu). Bu tuşların aynı çıktıyı üretmeleri fakat farklı konumdaki tuşlar olmaları sebebiyle alan tabanlı bir çalışma yapılmış ve bu sayede analiz sonuçlarının iyileştiği gözlemlenmiştir [7].

Uzun Y. yaptığı doktora tezi çalışmasında, önceden paylaşılmış erişme açık bir veri setini kullanarak YSA ile kimlik doğrulama analizleri yapmıştır ve en iyi sonuç olarak %7.73 EER oranı tespit edilmiştir. Ek olarak, geliştirilen veri toplama uygulamasıyla cinsiyet, yaş ve tuşlama dinamikleri toplanıp analiz edilerek, kimlik tespiti dışında yaş ve cinsiyet tahmini analizleri de yapılmıştır. Yaş tahmininde %8.20 EER oranı, cinsiyet

tahmininde %40 EER oranı tespit edildiği görülmektedir. Cinsiyet ayrıştırmasında başarılı gözükme de, yaş tahmininde önemli sonuçlar elde edilmiştir. Ayrıca çocuk kullanıcıların tespiti ve içerik filtreleme konularına da değinilmiştir [8].

1.3.4. Mobil Cihazlarda Tuşa Basış Dinamikleriyle Yapılan Kimlik Doğrulama Çalışmaları

Fridman L. yaptığı doktora tezi çalışmasında, hem masaüstü bilgisayarlar, hem de mobil cihazlar için davranışsal biyometrik kimlik doğrulama analizleri yapmıştır. Bilgisayarlar için yapılan çalışmada tuşa basma dinamikleri, fare hareketleri ve tıklamaları, metindeki sayı oranı, büyük harf oranı, en çok kullanılan karakterler gibi özelliklerle veri seti oluşturmuştur. Yapılan analizler sonucunda en iyi %1 Yanlış Kabul Oranı (FAR) ve %1 Yanlış Reddetme Oranı (FRR) oranları elde edilmiştir. Mobil cihazlar için yaptığı çalışmada ise tuşa basma dinamikleri, kullanılan uygulamalar, ziyaret edilen internet siteleri ve lokasyon verilerini kullanarak veri seti oluşturmuştur. Mobil cihazlar için yapılan analizler sonucunda en iyi %1 EER değerleri elde edilmiştir [9].

Çeker H. yaptığı doktora tezi çalışmasında, bilgisayarlar ve mobil cihazlar için toplanan, anonim veri setlerini kullanarak kimlik doğrulama sisteminin performansını yükseltmek için analizler yapmıştır. Performansı yükseltmek için bilgi aktarımı yöntemleri kullanılmıştır. Ayrıca literatürdeki ilk tuşa basış dinamikleri ve evrimsel sinir ağıları çalışmasını da gerçekleştirmiştir. Performans sıfıra yakın EER değerlerine kadar arttırılabilmektedir [10].

Muliono Y. ve arkadaşları konferans bildirilerinde, daha önce farklı çalışmalara konu olan, sadece tuşa basış dinamiklerini içeren ve önceden paylaşılmış halka açık olan veri setlerini, optimize edilmiş özel bir algoritmayla analiz etmişlerdir. Çalışmada EER değerlerine yer verilmemiş olup, en iyi sonuç olarak %92.60 doğruluk oranı tespit edilmiştir [11].

Krishnamoorthy S. yaptığı yüksek lisans tez çalışmasında, geliştirilmiş bir Android uygulaması aracılığıyla, katılımcılardan tuşa basış dinamikleri verileri, parmak alanı verileri, parmak basıncı verileri ve tuşa basarken ekranda hangi koordinatlara dokunulduğunu ifade eden verileri toplayarak kimlik doğrulaması analizi yapılmıştır. Geliştirilen uygulamanın içerisine butonlar vasıtasıyla tuşları anımsatan, gerçek sistem

klavyesi olmayan sanal bir klavye konumlandırılmıştır. Çalışma sonunda %97.40 doğruluk oranı elde edilmiştir [12]. Yapılan veri toplama işleminin tasarımı gereği, katılımcılar yüksek kullanım konforuna sahip mobil cihazların orijinal klavyelerini veya kendi tercih ettikleri klavyeleri kullanmamışlardır.

Syed Z.A. yaptığı doktora tezi çalışmasında, daha önceden paylaşılmış tuşa basış dinamikleri veri setleri ve çalışma kapsamında toplanmış tuşa basış dinamikleri veri setleriyle analizler yapmıştır. Kullanıcıların bildikleri ve sürekli yazdıkları metinlerdeki tuşa basma dinamiklerinin, ilk defa karşılaştıkları veya sıklıkla yazmadıkları metinlere göre daha kalıplaşmış ve kişiye özgü olduğu tespit edilmiştir. Ayrıca bilinen ve sık yazılan metinlerin yazma sürelerinin, bilinmeyen veya sıklıkla yazılmayan metinlere göre daha kısa olduğu tespit edilmiştir [13].

1.3.5. Mobil Cihazlarda Sensör Verileriyle Yapılan Kimlik Doğrulama Çalışmaları

Yoneda K. yaptığı yüksek lisans tezi çalışmasında, geliştirilen bir Android uygulaması vasıtasıyla katılımcılardan mobil cihazlar üzerlerindeyken günlük yaptıkları rutin hareketleri yapmalarını istemiş ve uygulama vasıtasıyla Akselerometre ve Jiroskop verilerini toplamıştır [14]. Çok geniş çaplı pozisyonlarda ve uzun süreli sensör verilerinin toplandığı bu çalışmada %99 doğruluk oranına yakın sonuçlarla kişilere kimlik doğrulama işlemi yapılabilmektedir. Yapılan bu çalışmadan anlaşıldığı üzere mobil cihazlarda bulunan sensörlerin, kişilerin tespit edilmesinde etkili olduğu görülmektedir.

Catal C. ve arkadaşları makalelerinde, kullanıcıların mobil cihazlarındaki akselerometre sensörlerinden gelen verileri toplayarak, hangi aktiviteyi yaptıklarını tahmin eden bir analiz yapmışlardır. Yüksek doğruluk oranlarında kullanıcıların hangi aktiviteyi gerçekleştirdiklerini tespit edebilmişlerdir [15].

Syed Z. ve arkadaşları makalelerinde, ekrana dokunma dinamikleri ile kullanıcıların duruş postürleri ve cihaz boyutunun dokunma dinamiklerini nasıl etkilediğini incelemişlerdir. Kullanıcıların duruş postürlerini akselerometre ve jiroskop sensörlerinden gelen veriler sayesinde sisteme aktarmışlardır [16]. Bu çalışma kapsamında dokunma dinamikleri kapsam dışı olsa da, kişilerin ayırt edilmesinde duruş postürlerinin; dolayısıyla akselerometre ve jiroskop verilerinin önemi vurgulanmıştır.

Maghsoudi J. yaptığı doktora tezi çalışmasında, kullanıcıların mobil cihazlarını, ona bakmak için yüzlerine doğru nasıl kaldırdıkları ve gelen çağrıları cevaplamak için nasıl kulaklarına götürdükleri hareketlerinden yola çıkarak; günlük hayatta herkes tarafından sık kullanılan iki aktivite seçmiştir. Seçtiği bu aktivitelere dair toplanan verileri Makine Öğrenmesi algoritmalarıyla analiz ederek, kimlik doğrulama analizleri yapıştır [17]. Çalışmada yapılan birden fazla analiz sonucunda ortalama %90 doğruluk oranı değerlerine ulaşmıştır.

Bhattarai A. yaptığı yüksek lisans tezi çalışmasında, mobil cihazlarda bulunan sensörler vasıtasıyla elde edilen veriler aracılığıyla kişilerin gün içinde yaptıkları hareketlerden yola çıkarak, sürekli doğrulama yapan bir sistemi analiz etmiş; en iyi EER oranını yürürken %2,11 olarak tespit etmiştir [18]. Çalışmadan da görüldüğü üzere sensör verileri kimlik doğrulama için anlamlı gözükmemektedir.

Mendizabal Vazquez I. ve arkadaşları konferans bildirilerinde, kullanıcılardan 4 haneli nümerik giriş metni toplayarak; tuşlama basıncı, parmak alanı ve akselerometre verilerini içeren bir veri seti üzerinden davranışsal biyometrik kimlik doğrulama analizi yapmıştır. En iyi EER oranı %20 olarak tespit edilmiştir. Bu şekilde yüksek bir hata oranının çıkmasının sebebi olarak düşük sayıda örnekleme çalışılması belirtilmiştir [19]. Bu tez çalışmasında, Mendizabal Vazquez I. ve arkadaşlarının konferans bildirilerinden farklı olarak; giriş metninin özelliklerinin sonuçları etkileyebileceği düşünülerek (bkz: Bölüm 2) 8 karakterli güçlü şifre paradigmasına uyan bir metin kullanılmıştır.

Yuksel A. ve arkadaşları makalelerinde, iPhone marka telefonlarda kullanılmak üzere geliştirdikleri uygulama vasıtasıyla; kullanıcılar yazı yazarken akselerometre ve jiroskop verilerini toplayıp kimlik doğrulama analizleri yapmışlardır. Yazılan metin ile sensör verilerini eşleştirmeden, sadece sensör verilerine odaklanmışlardır. Analizler sonucunda en iyi sonuç olarak %100 doğruluk oranı elde etmişlerdir [20]. Bu çalışma sayesinde iPhone markalı mobil cihazlarda sensör verileriyle kimlik doğrulama işleminin yapılabilirdiği görülmektedir.

Karakaya N. ve arkadaşları konferans bildirilerinde, daha önce farklı çalışmalara da konu olan; önceden paylaşılmış halka açık veri setlerindeki akselerometre, jiroskop ve manyetometre sensörlerinden toplanan verileri kullanarak davranışsal biyometrik kimlik

doğrulama analizi yapmışlardır. Çalışmada sadece sensör verileri konu edilmiş ve EER değerlerine yer verilmemiş olup, en iyi sonuç olarak % 99.60 doğruluk oranı verilmiştir [21].

Shen C. ve arkadaşları makalelerinde, mobil cihazlardaki akselerometre ve jiroskop sensörlerinden faydalanarak; kullanıcıların otururken, yürürken, tempolu yavaş koşarken ve zıplarken yaptıkları davranışları analiz etmişlerdir. Verileri toplarken mobil cihazları kullanıcıların ellerine, ceplerine ve kollarına yerleştirmişlerdir. En iyi sonuç olarak %2.21 EER değerini yürürken elde etmişlerdir [22].

Buriro A. ve arkadaşları makalelerinde, mobil cihazlarda bulunan akselerometre, jiroskop ve yerçekimi ölçer sensörlerini kullanarak; telefon çaldığı zaman kişilerin cevaplamak için ekranda parmaklarını kaydırmaları ve kulaklarına götürmeleri hareketlerinden kimlik doğrulama analizleri yapmışlardır. En iyi sonuç olarak %99.35 doğruluk oranıyla kimlik doğrulama işlemi yapılabilmektedir [23].

Lee W. yaptığı doktora tezi çalışmasında, mobil cihazlarda ve akıllı saatlerde bulunan sensörleri güvenlik bakış açısıyla değerlendirmiş, sensörlerin güvenlik açıklıklarına odaklanmıştır. Lee W.'nin doktora tezi çalışmasında sensörlerden veri toplanırken kullanıcıdan izin istenmediği belirtilmiş olup bu sensörlerin bilgi çıkarımı atakları için kullanılabileceğinden, sensör verilerinin kayıt altına alınıp, kullanıcı hareketlerinin taklit edilebileceğinden veya tekrar eğitimden geçirilerek modelin tespit edilebileceğinden söz edilmiştir. Ayrıca yapılan analizler sonucunda kullanıcı hakkında bilgi çıkarımı için en iyi iki sensörün akselerometre ve jiroskop sensörleri olduğu belirtilmiştir. Çalışmanın devamında, telefonu tutup kaldırma ve ele alma hareketinden güvenli kimlik doğrulama yapan bir çalışma yapılmış; en iyi sonuç % 96.30 doğruluk oranı, %0 FAR ve % 7.60 FRR olacak şekilde akselerometre ve jiroskop sensörlerinin beraber kullanıldığı veri setinden edinilmiştir. Ek olarak, mobil cihaz üzerinde bulunan sensörlerden gelen veriler kaydedilerek, kullanıcıların mobil cihazlarda elle yazı yazma esnasındaki yaptıkları hareketlerin haritasını çıkaran bir çalışma yapılmıştır. Kullanıcıların elle yazdıkları el yazısı harfleri % 94.40 doğruluk oranı ile tespit edilmiştir [24]. Bu tez çalışmasında tasarlanan sistemin (bkz: Bölüm 4), Lee W. tarafından yapılan çalışmada bahsedilen ataklara karşı güçlü ve zayıf yönleri tartışılmıştır.

1.3.6. Mobil Cihazlarda Tuşa Basış Dinamikleri ve Sensör Verileriyle Yapılan Kimlik Doğrulama Çalışmaları

Singh M. ve arkadaşları makalelerinde, tek bir özelliğe ait biyometrik verilerle kimlik doğrulama yöntemi yerine, birden fazla biyometrik özellikler kullanılarak daha iyi ve güvenli sistemler tasarlanabileceğinden bahsetmişlerdir. Yapılan araştırma çalışması temel olarak doğruluk değerlerinin artırılması ve yapılan atakların başka özellikte biyometrik verilerin desteği sayesinde engellenmesini konu edinmiştir [25]. Bu tez çalışmasında tuşa basış dinamikleri ve mobil cihazlardaki sensör verileri birleştirilerek analizler yapılmıştır. Sistemi atlatmaya çalışan kişilerin hem tuşa basış dinamiklerini hem de mobil sensör verilerini elde etmesi gerekmektedir. Bu bağlamda sadece bir veri tipine yapılan ataklara göre daha güvenlidir.

Lee H. ve arkadaşları makalelerinde, mobil cihazlarda tuşa basış dinamikleri verileri ve akselerometre, jiroskop ve rotasyon sensörlerinden gelen verileri kullanarak kimlik doğrulama analizleri yapmışlardır. Giriş metni olarak kullanıcılardan 6 haneli nümerik karakterler toplamışlar ve en iyi sonuç olarak %7.89 EER değerini elde etmişlerdir [26]. Bu tez çalışmasında ise Lee H. ve arkadaşlarının makalelerinde bahsedilen çalışmadan farklı olarak, davranışsal özelliklerin daha iyi ortaya çıkmasına yarayan, yaygın olarak kullanılan ve güçlü şifre paradigmasına uyan bir giriş metni aracılığıyla veriler toplanmış; bahsedildiği üzere (bkz: Bölüm 3) daha iyi sonuçlar elde edilmiştir.

Coakley M.J. yaptığı doktora tezi çalışmasında; tuşa basış dinamikleri, cihaz pozisyonu ve tuşa basarken ekranda hangi koordinatlara dokunulduğuna ait verileri içeren benzer bir çalışma yapmıştır. Ancak bu çalışmada toplanan metin sadece nümerik olarak seçilmiştir [27]. Çalışma sonundaki analiz sonrasında en iyi sonuç olarak %3.90 EER oranı tespit edilmiştir. Günümüzde, uygulama sağlayıcı firmalar tarafında siber güvenlik alanında farkındalık artmış ve birçok uygulama veya web sitesi, kullanıcılarını güçlü şifre kullanmaya doğrudan zorlamakta veya yönlendirmektedir. Bu tez çalışmasında ise Coakley M.J.'nin doktora tezinde bahsedilen çalışmadan farklı olarak, güçlü şifre özelliklerine sahip bir giriş metni ele alınmıştır ve detaylara yer verilmiştir (bkz: Bölüm 2). Bununla beraber bu tez çalışmasında, mobil cihaz klavyelerinde cihazların ekran boyutlarından kaynaklı tuşlama alanının dar bir alana sahip olması sebebiyle; sayılar, semboller ve harfler farklı sayfalarından seçilmiş, bir takım tuşlar vasıtasıyla bu sayfalar

arasında geiř yaptırılmıřtır. Bu sayfalar arasındaki geiřlerde yařanan gecikmeler ve cihaz pozisyonunda gözlemlenen deėiřikliklerin davranıřsal biyometrik analizin sonularını etkilemesi öngörölmüřtür.

Ahmadzadeh E. yaptıėı doktora tezi alıřmasında, bu tez alıřmasındaki benzer yaklařımlarla farklı bir yoldan analiz yapılmıř. Kullanıcılardan mobil cihazları vasıtasıyla tuřa basıř dinamikleri ve jiroskop verileri toplandıktan sonra, toplanan jiroskop verilerinden; kullanıcının oturma, yürüme ve ayakta durma pozisyonlarından hangisinde olduėu sınıflandırılmıřtır. Tuřa basıř dinamikleri bu sınıflandırma sonularıyla etiketlenerek Makine Öğrenme algoritmalarıyla analizler yapılmıřtır. En iyi FAR %1.40, en iyi FRR %5.3 olarak tespit edilmiř [28]. Bu tez alıřmasında ise bahsedildiėi üzere (bkz: Bölüm 2), akselerometre sensörü alıřmaya dahil edilerek, pozisyon sınıflandırması yapılmadan ham sensör verileriyle analiz yapılmıřtır.

1.3.7. Mobil Cihazlardaki Sensörler İin Yapılan Siber Tehdit alıřmaları

Wang L. yaptıėı yüksek lisans tezi alıřmasında, akselerometre, jiroskop ve ıřık sensörlerini kullanarak; kullanıcının hangi harflere bastıėına dair ıkarım yapmak üzere bir analiz yapmıřtır. Toplam 30 katılımcıdan 20700 adet tuřa basıř kaydedilmiřtir. Analizlerin sonucunda %48.20 doėruluk oranı elde edilmiřtir [29]. alıřmadan anlařıldıėı üzere, sensör verileriyle basılan harflerin tahmin edilmesinde düşük doėruluk oranı tespit edilmiř olsa bile, ileride yapılacak yeni alıřmalar sayesinde doėruluk oranı yükseltilebilir.

Yang Z. ve arkadařları konferans bildirilerinde, mobil cihazlarda yer alan sensörlerden gelen verileri analiz ederek; kullanıcıların ekranda hangi tuřa dokunduklarını tespit eden bir alıřma yapmıřlardır. En iyi %85 doėruluk oranı tespit edilmiřtir [30]. Mobil sensörlerin güvenliėi hakkındaki konulara ilgili kısımda (bkz: Bölüm 4) yer verilmiřtir.

2. MATERYAL VE YÖNTEM

Davranışsal biyometrik kimlik doğrulama analizlerinin yapılabilmesi ve veri setinin oluşturulabilmesi için geliştirilen uygulama vasıtasıyla gönüllü kullanıcılardan rızalarıyla birlikte verileri toplanmış, daha sonra Makine Öğrenmesi algoritmalarıyla analizler gerçekleştirilmiştir.

2.1. Tuşa Basış Dinamikleri ve Sensör Verileri

Bilgisayar ortamında yapılan tuşlama dinamiklerine dayalı çalışmalarda, bilgisayar kullanmanın doğası gereği genellikle klavye düz zemin üzerinde sabit bir şekilde durduğu için sadece tuşa basma dinamiklerini analiz etmek anlamlı gözükmemektedir. Mobil cihazları konu edinen bu tez çalışmasında ise tuşa basış dinamiklerine ek olarak, insanların mobilitesinden kaynaklı olarak cihazın anlık hareketlerinin ve insanların farklı duruş pozisyonlarının tuşa basış dinamiklerini etkilediği düşünülmüş, mobil cihazlarda bulunan akselerometre ve jiroskop sensörlerinden toplanan veriler de bu amaçla analize dahil edilmiştir.

Mobil cihazlarda, akselerometre ve jiroskop sensörleri dışında doğrusal ivme ölçer, dönme vektörü, adım sayar, cihaz oryantasyon sensörü gibi başka sensörler de mevcuttur. Bahsedilen diğer sensörler hem donanım tabanlı hem yazılım tabanlı olabilmektedirler. Eğer bu sensörlerden gelen veriler çalışmaya dahil edilirse, yazılım tabanlı oldukları takdirde elde edilen veriler, donanım tabanlı olan akselerometre ve jiroskop sensörlerinin verilerinden türetildiği için veri setinde birbirlerine bağlı, tekrar eden veriler bulundurulmuş olunur. Bu sebeple çalışmada kesinlikle donanım tabanlı olduğu bilinen akselerometre ve jiroskop sensörleri seçilmiştir [31].

2.1.1. Zaman

İnsanların yetenekleri ve yatkınlıklarından kaynaklı farklı hızlarda yazı yazdıkları gerçeğinden hareket edilerek, kullanıcıların bastıkları her harf arasındaki geçen süre toplanmıştır.

2.1.2. Akselerometre

Akselerometre sensörü mobil cihazın, yer çekimi kuvvetini hesaba katarak üç boyutlu uzaydaki X, Y ve Z eksenleri yönündeki hızlanma kuvvetini ölçmektedir. Bu sensörden gelen verilerin birimi m/s^2 'dir [31].

2.1.3. Jiroskop

Jiroskop sensörü mobil cihazın, üç boyutlu uzaydaki X, Y ve Z eksenleri etrafındaki dönme hızını ölçmektedir. Bu sensörden gelen verilerin birimi rad/s 'dir [31].

2.2. Veri ve Toplanması

Çalışma için geliştirilen uygulama vasıtasıyla gönüllü katılımcılardan daha önceden belirlenmiş, güçlü şifreyi temsil eden “**z?nC+3Av**” metni toplanmıştır. Güçlü şifreler tahmin edilmesi zor olan karmaşık şifrelerdir. Bu bağlamda kullanıcılardan toplanan metin aşağıda belirtilen özellikleri taşımaktadır [32];

- 8 harf
- Büyük harf
- Küçük harf
- Sayı
- Noktalama işareti

Aynı zamanda mobil cihazların ekran boyutlarının küçük olmasından kaynaklı olarak sayılar, harfler ve noktalama işaretleri klavyenin farklı sayfalarında olmaktadır. En doğal kullanıcı davranışlarını toplayabilmek adına tasarlanan giriş metni sayesinde kullanıcılar bu sayfalar arasında geçiş yapmaya zorlanmış, bu geçişlerin basılan harfler arasındaki süreyi etkilemesiyle veri setinin kişiye özgü olması özelliğinin artırılması hedeflenmiştir.

2.2.1. Veri Toplama Uygulaması

Kullanıcılardan davranışsal biyometrik verilerin toplanması amacıyla Android işletim sistemi üzerinde çalışan bir mobil uygulama geliştirilmiştir. Uygulamada bir hata yaşanması olasılığına karşı yedek olması açısından 2 adet güçlü şifreyi temsil eden metin toplanmıştır. Çalışmada 1 numaralı “**z?nC+3Av**” metni kullanılmıştır.

BBDCollector

Aşağıdaki alanları doldurun:

1) z?nC+3Av

Üstteki 1 numaralı metni giriniz..

2) T7S!o9.T

Üstteki 2 numaralı metni giriniz..

Bir rumuz bilgisi giriniz..

ÇALIŞTIR

SIFIRLA

X - acc:-0.35906672, gyro:-0.09435368, rot:0.015685536

Y - acc:5.558044, gyro:-0.010458086, rot:0.2850296

Z - acc:8.511443, gyro:-0.0038366932, rot:0.94914836

Şekil 1 Mobil Uygulama Veri Giriş Sayfası

BBDCollector

```
{'char':'z','time':2,'xacc':-0.5793238,'yacc':3.618824,'zacc':9.076449,'xgyr':-0.04792797,'ygyr':0.14409068,'zgyr':-0.024606096,'xrot':0.013088774,'yrot':0.17236039,'zrot':0.92892134}

{'char':'?', 'time':1902,'xacc':-1.2736124,'yacc':3.1160634,'zacc':9.155456,'xgyr':-0.1566619,'ygyr':0.0023700744,'zgyr':0.0144892465,'xrot':-0.009241428,'yrot':0.16908935,'zrot':0.9354225}

{'char':'n','time':918,'xacc':-1.0892668,'yacc':3.0274818,'zacc':9.057298,'xgyr':-0.043651916,'ygyr':-0.041001324,'zgyr':-0.0038366932,'xrot':-0.008626229,'yrot':0.16980949,'zrot':0.9351566}

{'char':'C','time':602,'xacc':-0.560171,'yacc':3.5470014,'zacc':9.126727,'xgyr':-0.17254438,'ygyr':0.005424395,'zgyr':0.031593457,'xrot':0.014611648,'yrot':0.16949725,'zrot':0.9321996}

{'char':'+', 'time':917,'xacc':-0.7373343,'yacc':2.8718653,'zacc':9.746798,'xgyr':-0.23546344,'ygyr':-0.09720088,'zgyr':-0.006280156,'xrot':-0.0027780186,'yrot':0.16392474,'zrot':0.9359658}

{'char':'3','time':295,'xacc':-0.26330277,'yacc':3.2597094,'zacc':9.4642935,'xgyr':-0.12367519,'ygyr':0.016419962,'zgyr':0.03831297,'xrot':
```

MAİL GÖNDER

CİHAZA KAYDET

Şekil 2 Mobil Uygulama Veri Görüntüleme ve Gönderme Sayfası

Uygulama çalışma prensibi olarak, her tuşa basıldığında zaman bilgisiyle akselerometre ve jiroskop sensörlerinin anlık değerlerini kaydetmektedir.

Gönüllülerden toplanan davranışsal biyometrik veriler, veri setinin hazırlanması evresinde programatik olarak ayrıştırılmasının kolay olması için JSON formatında hazırlanmıştır.

Gönüllü kullanıcılara, verileri topluca veya tek tek gönderebilme opsiyonu sunmak adına anlık olarak mail gönderme ve cihaz üzerine kaydetme seçenekleri sunulmuştur.

2.2.2. Veri Setinin Oluşturulması

Veri setleri oluşturulurken harfler, zaman verisi, akselerometre ve jiroskop sensörlerinin verileri için kısaltmalar belirlenmiştir.

Her harf için belirlenen kısaltmalar aşağıdadır.

Tablo 1 Girilen Metin Harf Kısaltmaları

Harf	Kısaltma
z	z
?	QM
n	n
C	C
+	PS
3	THREE
A	A
v	v

Her harfe basıldığı anda zaman bilgisi, akselerometre ve jiroskop sensörlerinden gelen veriler için belirlenen kısaltmalar aşağıdadır.

Tablo 2 Toplanan Verilerin Açıklamaları ve Kısaltmaları

Kısaltma	Açıklama
time	Bir önceki tuşa basılan zaman arasındaki fark.
xacc	Akselerometre sensöründen elde edilen X eksen değeri.
yacc	Akselerometre sensöründen elde edilen Y eksen değeri.
zacc	Akselerometre sensöründen elde edilen Z eksen değeri.
xgyr	Jiroskop sensöründen elde edilen X eksen değeri.
ygyr	Jiroskop sensöründen elde edilen Y eksen değeri.
zgyr	Jiroskop sensöründen elde edilen Z eksen değeri.

Makine öğrenmesi algoritmalarıyla analiz yapılabilmesi için katılımcılardan toplanan her örneklemin yer alabileceği, girilen karakterlerin (bkz: Tablo 1) ve toplanan verilerin (bkz: Tablo 2) kombinasyonundan temel veri seti oluşturulmuştur. Oluşturulan nihai veri setinin kolon başlıkları ve değer tipleri aşağıdaki tabloda belirtilmiştir.

Tablo 3 Veri Seti Parametreleri ve Değer Tipleri

Kolon Başlığı	Değer Tipi	Kolon Başlığı	Değer Tipi
z_time	nümerik	PS_time	nümerik
z_xacc	nümerik	PS_xacc	nümerik
z_yacc	nümerik	PS_yacc	nümerik
z_zacc	nümerik	PS_zacc	nümerik
z_xgyr	nümerik	PS_xgyr	nümerik
z_ygyr	nümerik	PS_ygyr	nümerik
z_zgyr	nümerik	PS_zgyr	nümerik
QM_time	nümerik	THREE_time	nümerik
QM_xacc	nümerik	THREE_xacc	nümerik
QM_yacc	nümerik	THREE_yacc	nümerik
QM_zacc	nümerik	THREE_zacc	nümerik
QM_xgyr	nümerik	THREE_xgyr	nümerik
QM_ygyr	nümerik	THREE_ygyr	nümerik
QM_zgyr	nümerik	THREE_zgyr	nümerik
n_time	nümerik	A_time	nümerik
n_xacc	nümerik	A_xacc	nümerik
n_yacc	nümerik	A_yacc	nümerik
n_zacc	nümerik	A_zacc	nümerik
n_xgyr	nümerik	A_xgyr	nümerik
n_ygyr	nümerik	A_ygyr	nümerik
n_zgyr	nümerik	A_zgyr	nümerik
C_time	nümerik	v_time	nümerik
C_xacc	nümerik	v_xacc	nümerik
C_yacc	nümerik	v_yacc	nümerik
C_zacc	nümerik	v_zacc	nümerik
C_xgyr	nümerik	v_xgyr	nümerik
C_ygyr	nümerik	v_ygyr	nümerik
C_zgyr	nümerik	v_zgyr	nümerik

Gönüllü kullanıcılardan toplanan veriler üzerinden (bkz: Tablo 3) ilişkili parametre/kolon başlıkları seçilerek programatik olarak 4 farklı veri seti oluşturulmuştur. Oluşturulan veri setlerinin kombinasyonları aşağıdadır;

- Zaman (8 parametre)
- Zaman + Akselerometre verileri (32 parametre)
- Zaman + Jiroskop verileri (32 parametre)
- Zaman + Akselerometre verileri + Jiroskop verileri (56 parametre)

Bu şekilde farklı veri setlerinin oluşturulmasının sebebi, tuşa basış dinamiklerine ek olarak incelenen akselerometre ve jiroskop sensörlerinin verilerinin doğruluk performansına olan faydasını veya zararını inceleyebilmektir. Sonuçlar kısmında bu konuya değinilmiştir fakat sonuçlara olan iyi veya kötü etkisi hariç tutularak mobil cihazlarda tuşa basış dinamikleri incelenirken, cihazın mobilitesinin hesaba katılmasının daha tutarlı bir yaklaşım olduğu düşünülmektedir.

2.2.3. Veri Setinin Demografik Yapısı

Veri setinde toplam 812 örneklem bulunmaktadır. Bu örneklemelerin 377 tanesi 1 kişiye, kalan 435 tanesi farklı 63 kişiye aittir. Veri setinin bu şekilde oluşturulmasındaki amaç; bir kişi için eğitilmiş ve oluşturulmuş davranışsal biyometrik kimlik doğrulama modelinin, fazla sayıda kişinin çok çeşitli davranışsal özellikleriyle olan farklılığını ortaya koyabilmektir. Başka bir deyişle, 1 kişinin doğru kişi olduğu, 63 kişinin saldırı odaklı oturum açma denemesinde bulunduğu şeklinde yorumlanabilir.

Ek olarak doğru ve yanlış ayırmada kullanılan sırasıyla 377 ve 435 örneklem sayısı dağılımı sayesinde, sınıflandırma için veri dengesi sağlanmıştır.

Tablo 4 Veri Seti Dağılımı

Kişi Sayısı	Örneklem Sayısı	Sınıflandırma Etiketi	Dağılım Oranı
1	377	Doğru	46.43%
63	435	Yanlış	53.57%

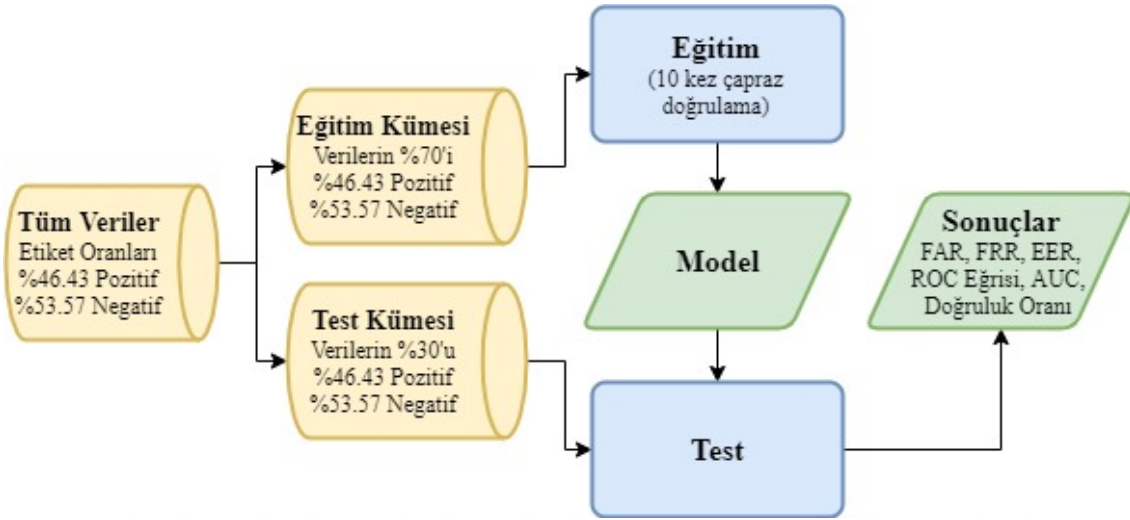
2.3. Verilerin Analizi

Bütün analizler R programlama dili kullanılarak yapılmıştır. R, istatistiksel hesaplama ve grafiksel gösterimler için kullanabilen ücretsiz bir programlama dilidir [33].

2.4. Analiz Yöntemi

Bu tez kapsamında 5 farklı makine öğrenmesi algoritması kullanılarak sınıflandırma yöntemiyle, oluşturulan 4 farklı veri seti analiz edilmiştir. Bu makine öğrenmesi algoritmalarına aşağıda (bkz: Bölüm 2.4.1) değinilmiştir.

Her algoritmayla yapılan analizlerde bir veri seti %70 eğitim ve %30 test gruplarına ayrılmış olup, grup içindeki verilerin doğru ve yanlış etiketlendirme oranları korunarak veriler rastgele seçilmiştir. Eğitim işlemleri 10 kez çapraz doğrulamaya tabi tutulmuştur.



Şekil 3 Analiz Yöntemi Akış Diagramı

Yapılan ön çalışmalarda, her analiz bir daha çalıştırıldığında rastgele seçimden kaynaklı eğitim ve test kümelerindeki veriler değiştiği için farklı sonuçlar gözlemlenmiştir. Bu nedenle, adil ve dengeli analiz sonuçları üretebilmek adına bir algoritma ve veri seti kombinasyonu için 5 kez analiz yapılmış; algoritma ve veri seti performanslarını kıyaslayabilmek için ortalamaya en yakın sonuçlar seçilmiş ve belirtilen (bkz: Bölüm 2.4.2) performans metriklerine göre değerlendirilmiştir.

2.4.1. Makine Öğrenmesi

Makine Öğrenmesi temel olarak, verileri toplamak ve ayrıştırmak, sonrasında geçmişe dönük oluşan veri setinden bir şeyler öğrenmek ve öğrenilen veriler üzerinden bir konu hakkında bir tespitte bulunmak veya tahmin yapmak için kullanılan yöntemleri çatısı altında barındıran genel bir kavramdır [34].

2.4.1.1. Karar Ağacı Sınıflandırıcı (Decision Tree Classifier)

Karar ağacı sınıflandırıcılar, veri setinde bulunan öznitelikler üzerinden basit kurallar çıkarıp, bu kuralları işleterek sonuçları tahmin etmeye yarayan bir Makine Öğrenmesi algoritmasıdır. Oluşturulan ağaç şeklindeki modellerde düğümler öznitelikleri, yapraklar ise sonuçları temsil eder [34].

2.4.1.2. Rastgele Orman Sınıflandırıcı (Random Forest Classifier)

Rastgele Orman algoritması, sınıflandırma sonucunu çok sayıda rastgele karar ağaçları üreterek, bu karar ağaçlarının ürettiği sonuçları birleştirip daha iyi hale getirmeye çalışan bir Makine Öğrenmesi algoritmasıdır [34].

2.4.1.3. Naive Bayes Sınıflandırıcı (Naive Bayes Classifier)

Bu algoritmanın çalışma şekli, bir veri için diğer tüm verilerin olasılığını hesaplayıp, olasılığın en yüksek olması durumuna göre sınıflandırma yapmasıdır. Algoritma Bayes Teoremi'ne dayalı çalışmaktadır. Bayes Teoremi 1812 yılında Thomas Bayes tarafından bulunmuştur [34].

2.4.1.4. Yapay Sinir Ağları (Neural Network)

Yapay sinir ağları (YSA) algoritmalarının amacı, çeşitli sonuçlara ve kararlara ulaşmak için insan beyninin düzenleme ve anlama şeklini taklit etmektir. Duyu organlarını kullanarak insan beyninin çalışması gibi YSA'da bilgiler veya değerler girdi katmanı tarafından toplanır, muhakeme yeteneği gibi gizli katman tarafından işlenir ve karara bağlanıp eylemde bulunulması gibi çıktı katmanı tarafından sonuçlar üretilir. Giriş ve çıkış katmanları basit girdilerden oluşmaktadırlar. Gizli katmanda ise, karmaşık işlemler yapabilen, birbirine bağlı birçok nöron mevcuttur. YSA'daki gizli katman veya katmanlar, veriler arasındaki ilişkileri öğrendikçe, nöronlar arasındaki bağlantılar için en iyi sonuçları verene kadar ayarlamalar yapar [34].

2.4.1.5. Yapay Sinir Ağları - Temel Bileşen Analizi (Neural Network – Principle Component Analysis)

Bu yöntemde (bkz: Bölüm 2.4.1.4) YSA'ya ek olarak Temel Bileşen Analizi yöntemi de kullanılmıştır. Çok sayıda özneliğin bulunduğu veri setlerinde, model oluşturulurken bazı öznitelikler yüksek öneme, bazıları ise düşük öneme sahip olmaktadır. Bu

bağlamda performansı iyileştirmek adına veri seti üzerinde yapılacak öznitelik ekleme ve çıkarma işlemleri zaman almaktadır. Ayrıca, düşük öneme sahip olan özniteliklerin veri setinden çıkartılması ise veri kaybına sebep olmaktadır. Bu yaklaşımda ise veri seti, en az veri kaybıyla anlamını kaybetmeden, korunarak daha az öznitelikli bulunduğu bir veri setine çevrilir.

2.4.2. Performans Göstergeleri

Tasarlanan davranışsal biyometrik kimlik doğrulama sisteminin performansı ölçülürken sadece doğruluk değeri değil, birden fazla parametre göz önünde bulundurularak optimum nokta elde edilmeye çalışılmıştır. Ölçüm metrikleri bu başlık altında detaylandırılmıştır.

2.4.2.1. Yanlış Reddetme Oranı (False Rejection Rate) (FRR)

Yanlış reddetme oranı, bir biyometrik kimlik doğrulama yöntemiyle yapılan doğrulama işleminin sonucunda, yanlışlıkla reddedilen işlemlerin oranını vermektedir. Biyometrik kimlik doğrulamayla oturum veya işlem izni verilmesi gereken gerçek kullanıcının, doğrulama işleminde başarısız olma durumudur. Gerçekleşen bu durumun oranının yüksek olması, biyometrik kimlik doğrulama sisteminin kullanıcı deneyimi açısından başarısız olması anlamına gelmektedir. Bu da, biyometrik kimlik doğrulama sisteminin entegre edildiği uygulamanın tercih edilebilirliğini azaltıcı bir faktördür.

Tablo 5 FAR Değerinin Tespiti

		Olması Gereken	
		Başarısız	Başarılı
Sistem	Başarısız	A	B
sonucu	Başarılı	C	D

$$(\%)FRR = \frac{\text{sistemin başarısız dediği, başarılı olması gerekenler}}{\text{başarılı olması gerekenlerin tamamı}} \times 100\%$$

$$(\%)FRR = \frac{B}{B + D} \times 100\%$$

2.4.2.2. Yanlış Kabul Oranı (False Acceptance Rate) (FAR)

Yanlış kabul oranı, bir biyometrik kimlik doğrulama yöntemiyle yapılan doğrulama işleminin sonucunda, yanlışlıkla doğrulanan işlemlerin oranını vermektedir. Biyometrik kimlik doğrulamayla oturum veya işlem izni verilmemesi gereken yetkisiz kullanıcıların, doğrulama işleminde başarılı olma durumudur. Gerçekleşen bu durumun oranının yüksek olması, biyometrik kimlik doğrulama sisteminin güvenlik açısından yetersiz, başka bir deyişle başarısız olması anlamına gelmektedir. Bu da, biyometrik kimlik doğrulama sisteminin entegre edildiği uygulamanın tercih edilebilirliğini azaltıcı bir diğer faktördür.

Tablo 6 FRR Değerinin Tespiti

		Olması Gereken	
		Başarısız	Başarılı
Sistem	Başarısız	A	B
sonucu	Başarılı	C	D

$$(\%)FAR = \frac{\text{sistemin başarılı dediği, başarısız olması gerekenler}}{\text{başarısız olması gerekenlerin tamamı}} \times 100\%$$

$$(\%)FAR = \frac{C}{A + C} \times 100\%$$

2.4.2.3. Alıcı İşletim Karakteristiği (Receiver Operating Characteristic) (ROC) Eğrisi ve Eğri Altındaki Alan (Area Under Curve)(AUC)

ROC eğrisi yapılan tahmin analizinin doğruyu ve yanlış ayırmada ne kadar güvenilebilir olduğu ölçmeye yarayan bir eğridir. Bu eğrinin altında kalan alan yani AUC metriği bize doğrudan, yapılan diğer analizlerle karşılaştırılabilir ve analizin başarısını ölçen sayısal veriyi vermektedir.

ROC eğrisi, Doğru Pozitif Oranı (True Positive Rate) (TPR) ve Yanlış Pozitif Oranı (False Positive Rate) (FPR) metrikleri koordinat düzleminde sırasıyla X ve Y eksenlerine koyularak çizilmektedir.

Bu çalışma kapsamında TPR ve FPR değerlerinin hesaplanması aşağıda belirtilmiştir;

Tablo 7 ROC Eğrisinin Tespiti

		Olması Gereken	
		Başarısız	Başarılı
Sistem	Başarısız	A	B
sonucu	Başarılı	C	D

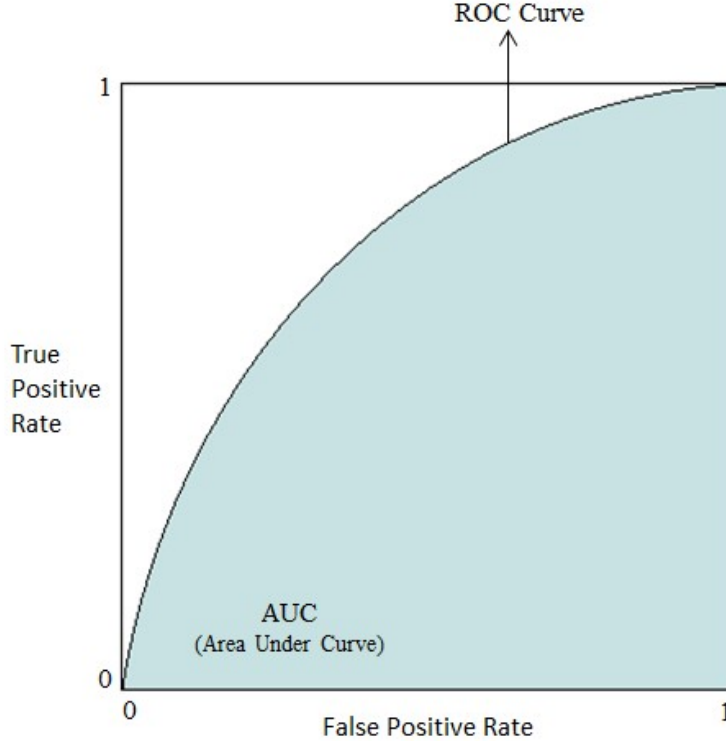
$$(\%)TPR = \frac{\text{sistemin başarılı dediği, başarılı olması gerekenler}}{\text{başarılı olması gerekenlerin tamamı}} \times 100\%$$

$$(\%)TPR = \frac{D}{B + D} \times 100\%$$

$$(\%)FPR = \frac{\text{sistemin başarılı dediği, başarısız olması gerekenler}}{\text{başarısız olması gerekenlerin tamamı}} \times 100\%$$

$$(\%)FPR = \frac{C}{A + C} \times 100\%$$

Bu bağlamda FPR değeri aslında daha önce bahsedilen (bkz: Bölüm 2.4.2.2) FAR değeri ile aynıdır.

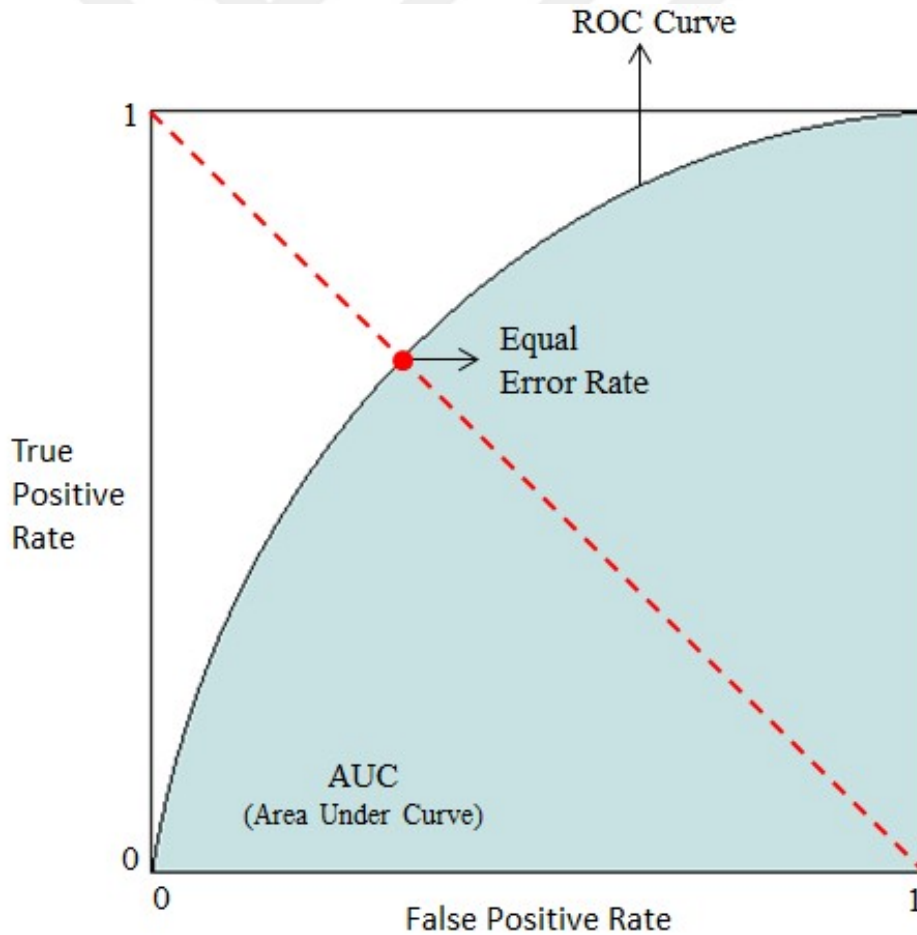


Şekil 4 Örnek ROC Eğrisi

2.4.2.4. Eş Hata Oranı (Equal Error Rate) (EER)

Biyometrik kimlik doğrulama çözümünün, FAR metriğinin yüksek olmasından kaynaklı güvenliğinin zayıf; FRR metriğinin yüksek olmasından kaynaklı kullanıcı deneyiminin kötü olduğundan ilgili bölümde (bkz: Bölüm 2.4.2.1 ve 2.4.2.2) bahsedilmektedir. Biyometrik kimlik doğrulama analizlerinin performansının, güvenlik ve kullanıcı deneyimi kavramlarından (FAR, FRR) ödün vermeden en dengeli ve doğru şekilde tutularak sistemin eşit hata oranı Equal Error Rate (EER) metriği ile ölçülmesi beklenir [1].

EER metriği FAR ve FRR oranlarının eşit olduğu noktayı temsil eder. Bu sayede kullanıcı deneyimi ve güvenlik arasında denge sağlanmaktadır. En başarılı sistem EER değerinin sayısal olarak en küçük olduğu sistemdir. EER metriği aşağıdaki örnek ROC eğrisi üzerinde işaretlenmiştir.



Şekil 5 Örnek ROC Eğrisi Üzerinde EER Noktası

2.4.2.5. Doğruluk Oranı (Accuracy)

Aşağıdaki formüller çalışma kapsamında yapılan analizlerin doğruluk oranı vermektedir. Formüller Tablo 7 referans alınarak yazılmıştır.

$$(\%) \text{Doğruluk Oranı} = \frac{\text{sistemin doğru tahmin ettikleri}}{\text{sistemin yaptığı tüm tahminler}} \times 100\%$$

$$(\%) \text{Doğruluk Oranı} = \frac{A + D}{A + B + C + D} \times 100\%$$

3. BULGULAR VE TARTIŞMA

Bu bölüm altında çalışmada bahsedilen (bkz: Bölüm 2) Makine Öğrenmesi algoritmalarıyla ve farklı veri seti kombinasyonlarıyla yapılan analizlerin sonuçları paylaşılmıştır. Analizler sonucunda ölçüm ve performans metrikleri olan Doğruluk Oranı, FAR, FRR, EER ve AUC değerleri kaydedilmiş, ROC eğrileri grafiksel olarak gösterilmiştir. Ayrıca sonuçların kolay yorumlanabilmesi için tüm analizlerin ve ortalama analizlerin topluca gösterildiği tablolar paylaşılmıştır.

3.1. Tüm Sonuçlar

Bu bölümde 4 farklı veri seti ve 5 farklı Makine Öğrenmesi algoritması kullanılarak, bunların her bir kombinasyonu için yapılan 5 analizin, toplamda 100 adet olmak üzere sonuçları paylaşılmıştır.

Bu çalışmadaki değerlendirmeler (bkz: Bölüm 4) adil ve dengeli bir kıyaslama için (bkz: Bölüm 2) ortalama değerler üzerinden yapılacaktır. Tüm analizlerin sonuçları, hesaplanan ortalama EER değerleri ile birlikte aşağıdaki tabloda paylaşılmış, ortalamaya en yakın analiz sonuçları tablo üzerindeki kalın olarak yazılmış satırlarla vurgulanmıştır.

Tablo 8 Tüm Analiz Sonuçları

Test No	Algoritma	Veri Seti	Doğruluk	FAR	FRR	EER	AUC	Ort. EER
1	Rastgele Orman	Zaman	99.16%	0.00%	1.77%	1.50%	0.999	1.94%
2			98.32%	0.80%	2.65%	2.30%	0.996	
3			97.90%	1.60%	2.65%	2.10%	0.997	
4			98.74%	0.00%	2.65%	1.50%	0.999	
5			97.48%	1.60%	3.54%	2.30%	0.998	
1	Rastgele Orman	Zaman+Aks.	97.90%	1.60%	2.65%	2.10%	0.999	1.72%
2			98.32%	2.40%	0.88%	1.80%	0.999	
3			98.74%	0.80%	1.77%	1.50%	0.999	
4			97.06%	1.60%	4.42%	1.70%	0.999	
5			98.32%	0.80%	2.65%	1.50%	0.999	
1	Rastgele Orman	Zaman+Jir	97.32%	1.80%	3.54%	3.00%	0.999	1.98%
2			97.77%	1.80%	2.65%	2.50%	0.998	
3			97.32%	0.90%	4.42%	1.00%	0.999	

4			98.21%	0.90%	2.65%	1.90%	1	
5			98.66%	0.00%	2.65%	1.50%	0.999	
1	Rastgele Orman	Zaman+Aks.+Jir.	97.32%	0.90%	4.42%	0.90%	0.999	
2			97.77%	2.70%	1.77%	1.50%	0.998	
3			97.32%	3.60%	1.77%	1.60%	0.998	1.34%
4			98.66%	1.80%	0.88%	1.30%	1	
5			98.21%	2.70%	0.88%	1.40%	1	
1	Naive Bayes	Zaman	97.06%	0.80%	5.04%	2.40%	0.986	
2			94.54%	0.80%	10.62%	4.55%	0.974	
3			95.80%	0.00%	8.85%	3.80%	0.985	4.01%
4			92.86%	2.40%	12.39%	5.10%	0.977	
5			94.96%	1.60%	8.85%	4.20%	0.976	
1	Naive Bayes	Zaman+Aks.	90.76%	11.20%	7.08%	7.50%	0.971	
2			96.22%	3.20%	4.42%	3.11%	0.994	
3			93.28%	6.40%	7.08%	6.60%	0.982	5.40%
4			91.20%	5.60%	6.59%	5.90%	0.991	
5			95.38%	5.60%	5.36%	3.90%	0.993	
1	Naive Bayes	Zaman+Jir.	98.66%	0.90%	1.77%	1.40%	0.999	
2			95.09%	5.41%	4.42%	4.50%	0.978	
3			96.43%	2.70%	4.42%	2.60%	0.988	2.64%
4			98.66%	0.00%	2.65%	2.20%	0.986	
5			97.32%	0.90%	4.42%	2.50%	0.99	
1	Naive Bayes	Zaman+Aks.+Jir.	94.64%	2.70%	7.96%	5.10%	0.989	
2			96.88%	0.00%	6.19%	3.30%	0.996	
3			98.21%	0.90%	2.65%	1.70%	0.999	2.86%
4			95.98%	3.60%	4.42%	3.50%	0.995	
5			99.11%	1.80%	0.00%	0.70%	1	
1	YSA	Zaman	94.54%	2.40%	8.85%	6.80%	0.97	
2			94.12%	4.00%	7.96%	6.00%	0.975	
3			89.92%	8.00%	12.39%	12.00%	0.897	7.52%
4			93.70%	1.60%	11.50%	7.50%	0.985	
5			95.80%	5.60%	2.65%	5.30%	0.983	
1	YSA	Zaman+Aks.	88.66%	13.60%	8.85%	13.40%	0.904	
2			91.60%	3.20%	14.16%	7.60%	0.975	
3			92.86%	4.00%	10.62%	8.40%	0.968	9.06%
4			90.76%	12.00%	6.19%	10.70%	0.929	

5			94.54%	6.40%	4.42%	5.20%	0.988	
1			93.75%	3.60%	8.85%	7.60%	0.95	
2			92.86%	5.41%	8.85%	6.90%	0.956	
3		YSA	94.20%	4.50%	7.08%	5.70%	0.972	7.44%
4			90.62%	6.31%	12.39%	8.50%	0.967	
5		Zaman+Jir.	91.96%	9.91%	6.19%	8.50%	0.951	
1			93.30%	3.60%	9.73%	7.20%	0.972	
2			91.96%	9.01%	7.08%	8.00%	0.941	
3		YSA	91.52%	10.81%	6.19%	9.50%	0.956	8.38%
4			90.62%	9.91%	8.85%	9.60%	0.935	
5		Zaman+Aks.+Jir.	91.07%	12.61%	5.31%	7.60%	0.954	
1			98.74%	1.60%	0.88%	1.60%	0.997	
2			97.06%	2.40%	3.54%	2.40%	0.99	
3		YSA-PCA	97.48%	4.00%	0.88%	2.35%	0.995	2.55%
4			97.06%	1.60%	4.42%	2.40%	0.999	
5		Zaman	96.22%	1.60%	6.19%	4.00%	0.996	
1			96.22%	2.40%	5.31%	3.30%	0.994	
2			97.48%	2.40%	2.65%	2.60%	0.998	
3		YSA-PCA	98.74%	1.60%	0.88%	0.80%	1	2.12%
4			98.74%	0.80%	1.77%	1.50%	0.999	
5		Zaman+Aks.	96.22%	4.80%	2.65%	2.40%	0.997	
1			97.77%	1.80%	2.65%	2.10%	0.995	
2			97.32%	4.50%	0.88%	2.10%	0.998	
3		YSA-PCA	97.77%	2.70%	1.77%	2.20%	0.989	2.38%
4			97.32%	3.60%	1.77%	2.10%	0.999	
5		Zaman+Jir.	97.77%	4.50%	0.00%	3.40%	0.989	
1			98.66%	0.90%	1.77%	1.50%	0.990	
2			99.11%	1.80%	0.00%	0.85%	0.999	
3		YSA-PCA	98.21%	0.90%	2.65%	2.30%	0.998	1.87%
4			96.43%	3.60%	3.54%	3.20%	0.989	
5		Zaman+Aks.+Jir.	99.11%	0.00%	1.77%	1.50%	0.995	
1			92.02%	3.20%	13.27%	12.00%	0.926	
2			91.60%	11.20%	5.31%	9.00%	0.938	
3		Karar Ağacı	88.66%	11.20%	11.50%	11.50%	0.898	11.16%
4			85.71%	13.60%	15.04%	14.80%	0.879	
5		Zaman	92.44%	6.40%	8.85%	8.50%	0.932	

1	Karar Ağacı	Zaman+Aks.	87.82%	15.20%	8.85%	13.00%	0.894	11.54%
2			92.02%	10.40%	5.31%	9.50%	0.938	
3			89.08%	8.00%	14.16%	11.20%	0.907	
4			86.13%	10.40%	17.70%	14.00%	0.884	
5			91.18%	7.20%	10.62%	10.00%	0.913	
1	Karar Ağacı	Zaman+Jir.	86.16%	15.32%	12.39%	14.00%	0.897	10.62%
2			89.29%	8.11%	13.27%	12.50%	0.902	
3			92.86%	6.31%	7.96%	7.70%	0.933	
4			91.52%	5.41%	11.50%	8.70%	0.929	
5			90.62%	10.81%	7.96%	10.20%	0.921	
1	Karar Ağacı	Zaman+Aks.+Jir.	89.29%	12.07%	8.85%	11.20%	0.924	11.56%
2			91.07%	11.71%	6.19%	10.00%	0.934	
3			91.96%	5.41%	10.62%	10.00%	0.924	
4			87.50%	7.21%	17.70%	15.60%	0.883	
5			90.18%	11.71%	7.96%	11.00%	0.911	

3.2. Ortalama Analiz Sonuçları

3.2.1. Karar Ağacı Sınıflandırıcı Çalışmaları

Tüm Karar Ağacı Sınıflandırıcı çalışmaları R dili üzerinde CARET kütüphanesi ve CART algoritmasını kullanan “rpart” methodu kullanılarak yapılmıştır [35].

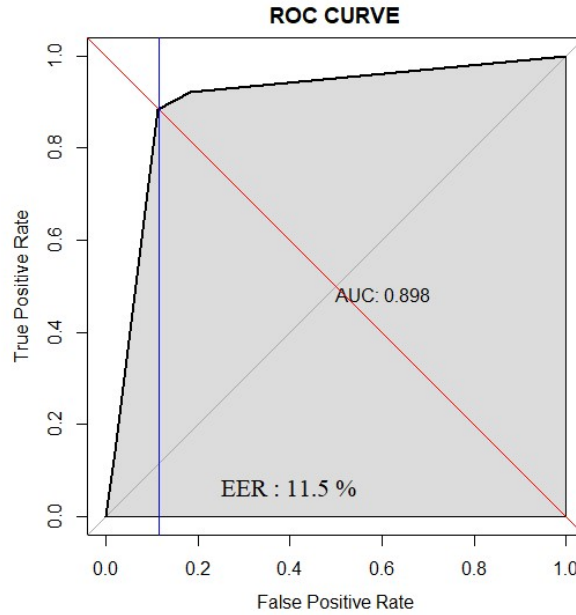
3.2.1.1. Zaman Veri Seti

Karar Ağacı Sınıflandırıcı yöntemiyle Zaman veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 9 Karar Ağaçları Algoritmasıyla Zaman Veri Setinin Sonuçları

Algoritma	Decision Tree
FAR	11.20%
FRR	11.15%
EER	11.50%
AUC	0.898
Doğruluk	88.66%
Veri Seti	Zaman

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 6 Karar Ağaçları Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası

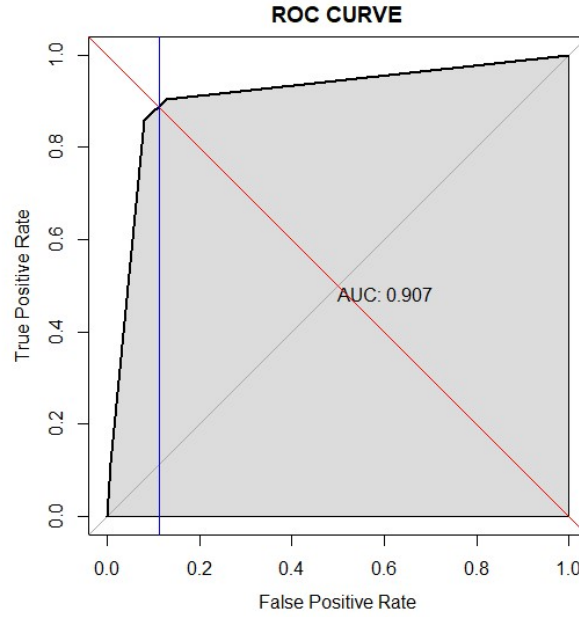
3.2.1.2. Zaman ve Akselerometre Veri Seti

Karar Ağacı Sınıflandırıcı yöntemiyle Zaman ve Akselerometre veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 10 Karar Ağaçları Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları

Algoritma	Decision Tree
FAR	8.00%
FRR	14.16%
EER	11.20%
AUC	0.907
Doğruluk	89.08%
Veri Seti	Zaman ve Akselerometre

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 7 Karar Ağaçları Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası

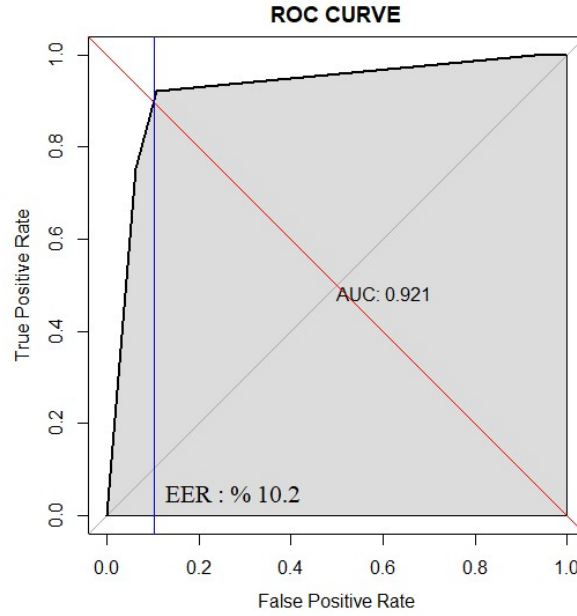
3.2.1.3. Zaman ve Jiroskop Veri Seti

Karar ağacı sınıflandırıcı yöntemiyle Zaman ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 11 Karar Ağaçları Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları

Algoritma	Decision Tree
FAR	10.81%
FRR	7.96%
EER	10.20%
AUC	0.921
Doğruluk	90.62%
Veri Seti	Zaman ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 8 Karar Ağaçları Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

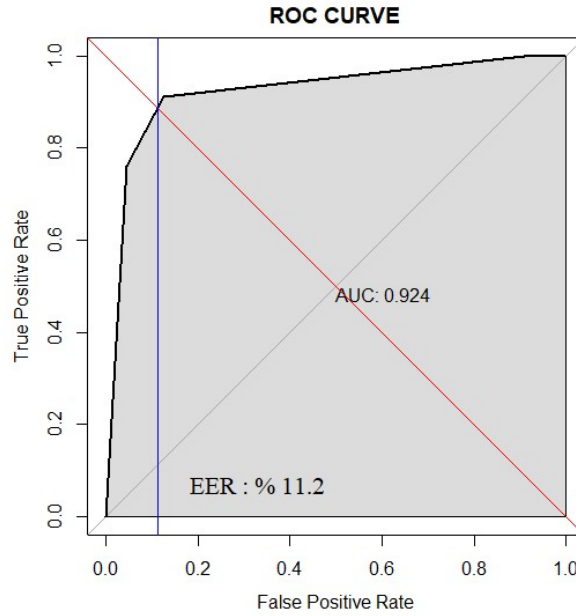
3.2.1.4. Zaman, Akselerometre ve Jiroskop Veri Seti

Karar ağacı sınıflandırıcı yöntemiyle Zaman, Akselerometre ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 12 Karar Ağaçları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları

Algoritma	Decision Tree
FAR	12.07%
FRR	8.85%
EER	11.20%
AUC	0.924
Doğruluk	89.26%
Veri Seti	Zaman, Akselerometre ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 9 Karar Ağaçları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

3.2.2. Rastgele Orman Sınıflandırıcı Çalışmaları

Tüm Rastgele Orman Sınıflandırıcı çalışmaları R dili üzerinde CARET kütüphanesi ve Rastgele Orman algoritmasını kullanan “rf” methodu kullanılarak yapılmıştır [35].

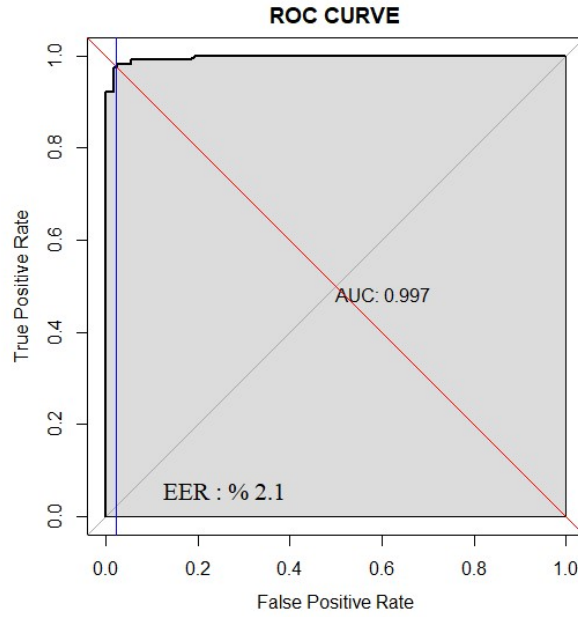
3.2.2.1. Zaman Veri Seti

Rastgele Orman Sınıflandırıcı yöntemiyle Zaman veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 13 Rastgele Ormanlar Algoritmasıyla Zaman Veri Setinin Sonuçları

Algoritma	Random Forest
FAR	1.60%
FRR	2.65%
EER	2.10%
AUC	0.997
Doğruluk	97.90%
Veri Seti	Zaman

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 10 Rastgele Ormanlar Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası

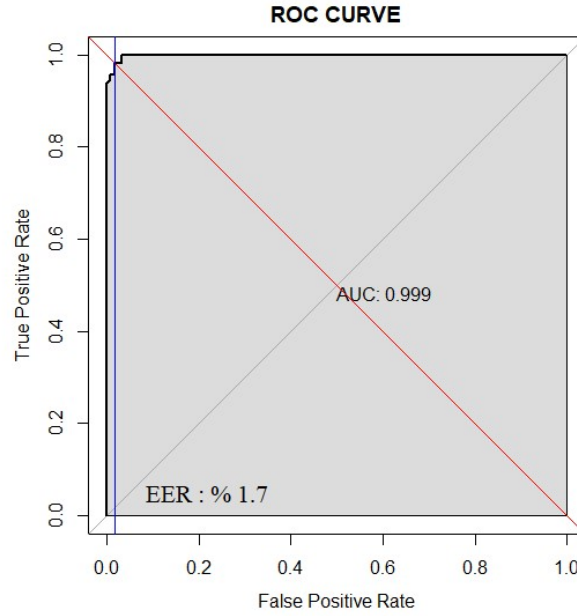
3.2.2.2. Zaman ve Akselerometre Veri Seti

Rastgele Orman Sınıflandırıcı yöntemiyle Zaman ve Akselerometre veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 14 Rastgele Ormanlar Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları

Algoritma	Random Forest
FAR	1.60%
FRR	4.42%
EER	1.70%
AUC	0.999
Doğruluk	97.06%
Veri Seti	Zaman ve Akselerometre

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 11 Rastgele Ormanlar Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası

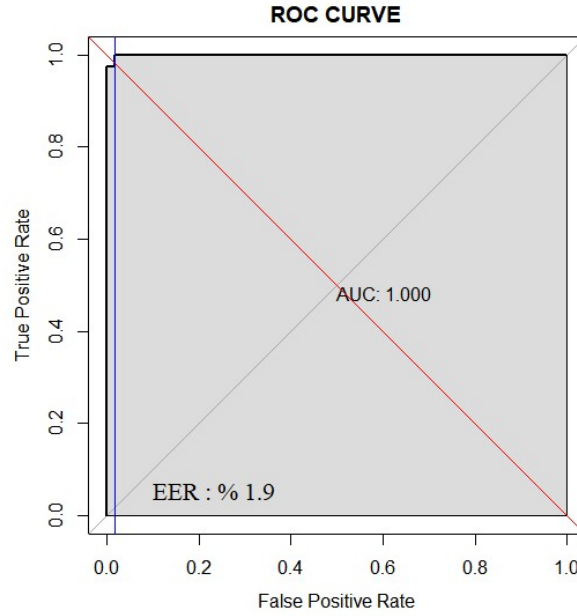
3.2.2.3. Zaman ve Jiroskop Veri Seti

Rastgele Orman Sınıflandırıcı yöntemiyle Zaman ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 15 Rastgele Ormanlar Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları

Algoritma	Random Forest
FAR	0.90%
FRR	2.65%
EER	1.90%
AUC	1
Doğruluk	98.21%
Veri Seti	Zaman ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 12 Rastgele Ormanlar Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

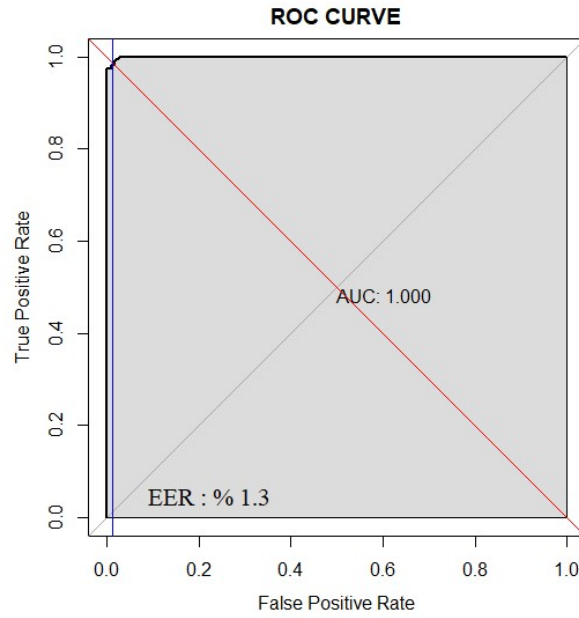
3.2.2.4. Zaman, Akselerometre ve Jiroskop Veri Seti

Rastgele Orman Sınıflandırıcı yöntemiyle Zaman, Akselerometre ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 16 Rastgele Ormanlar Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları

Algoritma	Random Forest
FAR	1.80%
FRR	0.88%
EER	1.30%
AUC	1
Doğruluk	98.66%
Veri Seti	Zaman, Akselerometre ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 13 Rastgele Ormanlar Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

3.2.3. Naive Bayes Sınıflandırıcı Çalışmaları

Bu algoritmayla yapılan tüm çalışmalar Bayes Teoremine dayandırılan Naive Bayes Sınıflandırıcı ile yapılmıştır. Tüm Naive Bayes Sınıflandırıcı çalışmaları R dili üzerinde CARET kütüphanesi ve “nb” methodu kullanılarak yapılmıştır [35].

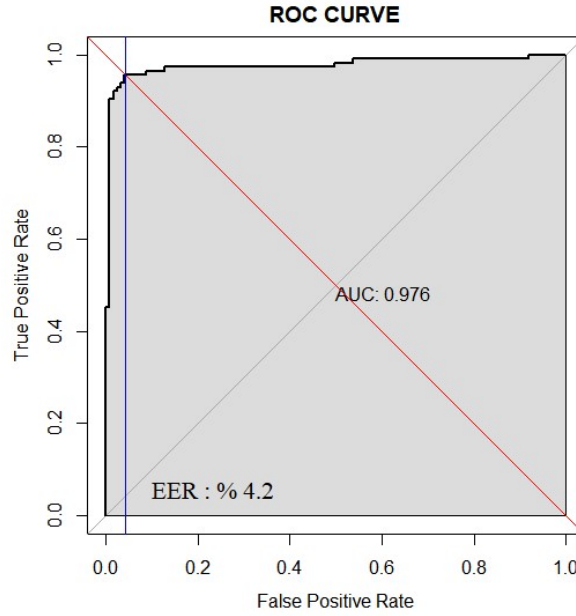
3.2.3.1. Zaman Veri Seti

Naive Bayes Sınıflandırıcı yöntemiyle Zaman veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 17 Naive Bayes Algoritmasıyla Zaman Veri Setinin Sonuçları

Algoritma	Naive Bayes
FAR	1.60%
FRR	8.85%
EER	4.20%
AUC	0.976
Doğruluk	94.96%
Veri Seti	Zaman

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 14 Naive Bayes Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası

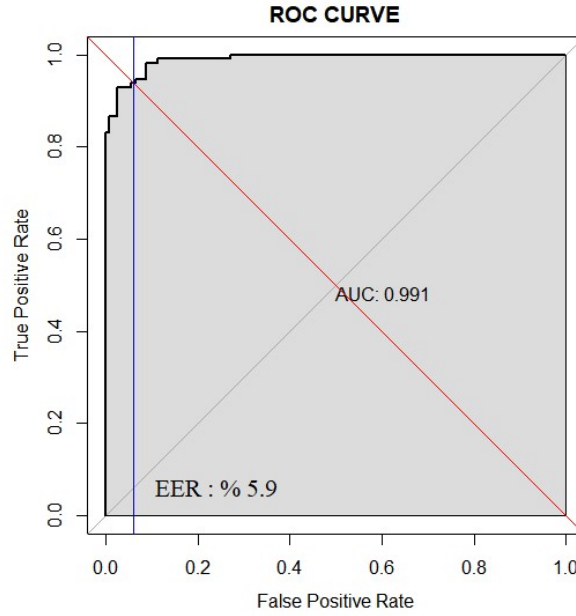
3.2.3.2. Zaman ve Akselerometre Veri Seti

Naive Bayes Sınıflandırıcı yöntemiyle Zaman ve Akselerometre veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 18 Naive Bayes Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları

Algoritma	Naive Bayes
FAR	5.60%
FRR	6.59%
EER	5.90%
AUC	0.991
Doğruluk	91.20%
Veri Seti	Zaman ve Akselerometre

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 15 Naive Bayes Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası

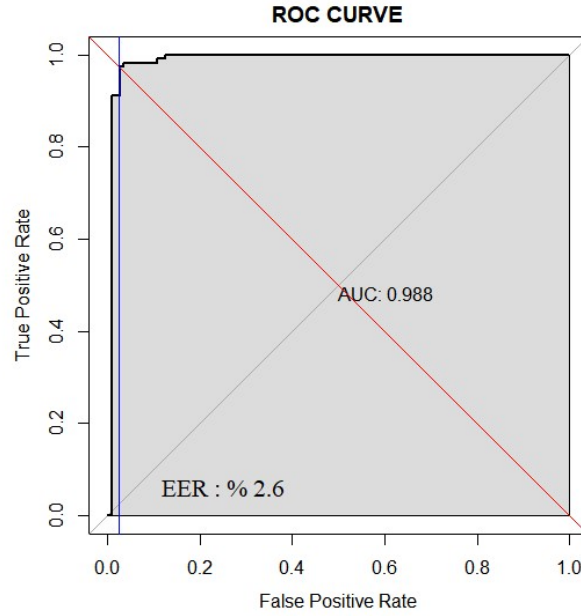
3.2.3.3. Zaman ve Jiroskop Veri Seti

Naive Bayes Sınıflandırıcı yöntemiyle Zaman ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 19 Naive Bayes Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları

Algoritma	Naive Bayes
FAR	2.70%
FRR	4.42%
EER	2.60%
AUC	0.988
Doğruluk	96.43%
Veri Seti	Zaman ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 16 Naive Bayes Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

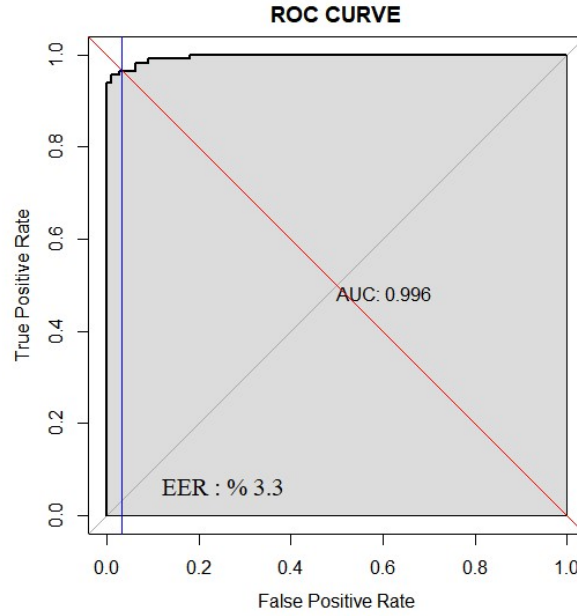
3.2.3.4. Zaman, Akselerometre ve Jiroskop Veri Seti

Naive Bayes Sınıflandırıcı yöntemiyle Zaman, Akselerometre ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 20 Naive Bayes Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları

Algoritma	Naive Bayes
FAR	0.00%
FRR	6.19%
EER	3.30%
AUC	0.996
Doğruluk	96.88%
Veri Seti	Zaman, Akselerometre ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 17 Naive Bayes Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

3.2.4. Yapay Sinir Ağları Çalışmaları

Bu algoritmayla yapılan tüm çalışmalarda ileri beslemeli ve tek gizli katmanlı YSA kullanılmıştır. Tüm YSA çalışmaları R dili üzerinde CARET kütüphanesi ve “nnet” methodu kullanılarak yapılmıştır [35].

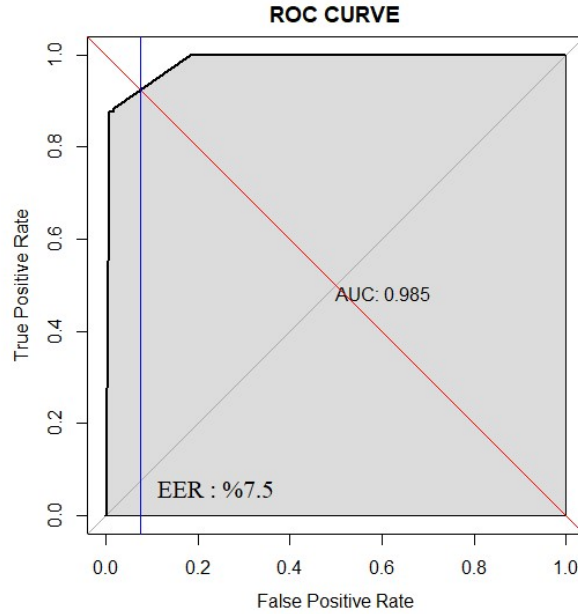
3.2.4.1. Zaman Veri Seti

YSA yöntemiyle Zaman veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 21 Yapay Sinir Ağları Algoritmasıyla Zaman Veri Setinin Sonuçları

Algoritma	Neural Network
FAR	1.60%
FRR	11.50%
EER	7.50%
AUC	0.985
Doğruluk	93.70%
Veri Seti	Zaman

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 18 Yapay Sinir Ağları Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası

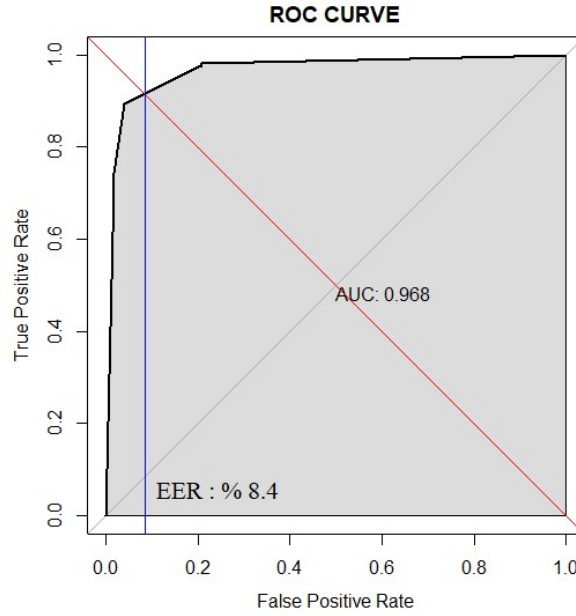
3.2.4.2. Zaman ve Akselerometre Veri Seti

YSA yöntemiyle Zaman ve Akselerometre veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 22 Yapay Sinir Ağları Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları

Algoritma	Neural Network
FAR	4.00%
FRR	10.62%
EER	8.40%
AUC	0.968
Doğruluk	92.86%
Veri Seti	Zaman ve Akselerometre

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 19 Yapay Sinir Ağları Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası

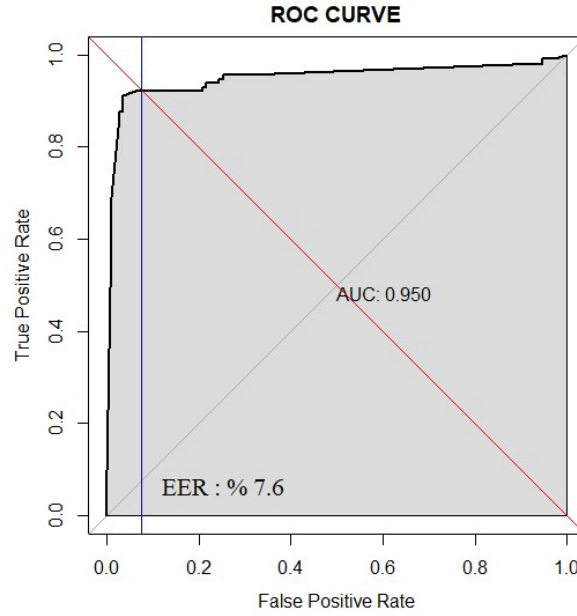
3.2.4.3. Zaman ve Jiroskop Veri Seti

YSA yöntemiyle Zaman ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 23 Yapay Sinir Ağları Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları

Algoritma	Neural Network
FAR	3.60%
FRR	8.85%
EER	7.60%
AUC	0.95
Doğruluk	93.75%
Veri Seti	Zaman ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 20 Yapay Sinir Ağları Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

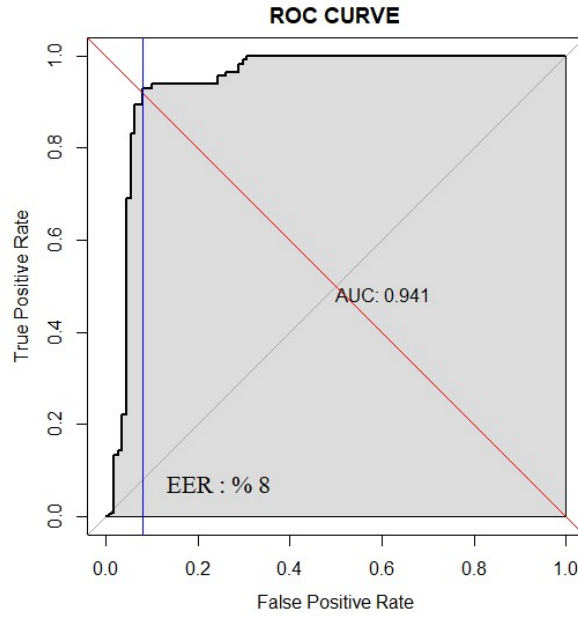
3.2.4.4. Zaman, Akselerometre ve Jiroskop Veri Seti

YSA yöntemiyle Zaman, Akselerometre ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 24 Yapay Sinir Ağları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları

Algoritma	Neural Network
FAR	9.01%
FRR	7.08%
EER	8.00%
AUC	0.941
Doğruluk	91.96%
Veri Seti	Zaman, Akselerometre ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.

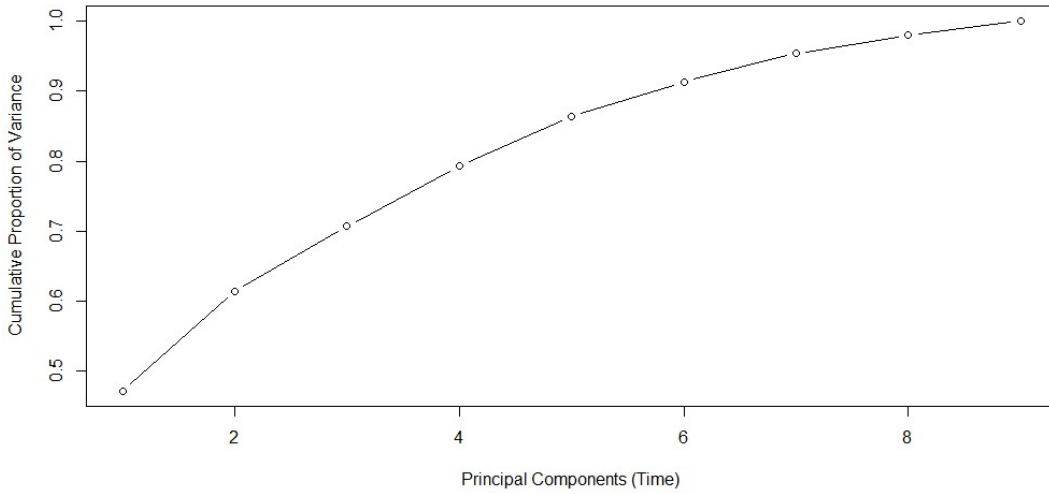


Şekil 21 Yapay Sinir Ağları Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

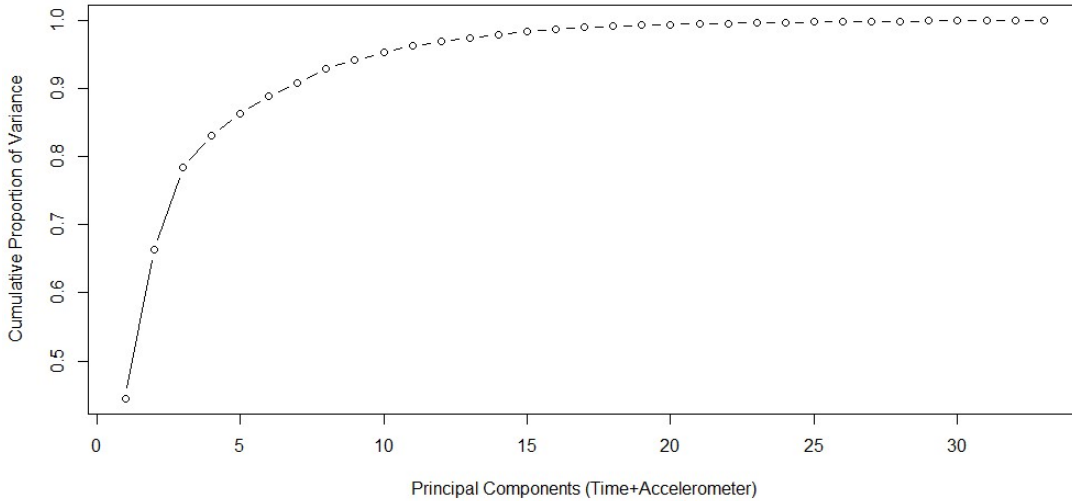
3.2.5. Yapay Sinir Ağları - Temel Bileşen Analizi Çalışmaları

Bu algoritmayla yapılan tüm çalışmalarda YSA Çalışmaları'nda (bkz: Bölüm 3.2.4) kullanılan yöntem ek olarak temel bileşen analizi yöntemi kullanılmıştır. Tüm YSA çalışmaları R dili üzerinde bulunan CARET kütüphanesi, “nnet” methodu ve “preProcOptions” parametresi kullanılarak yapılmıştır [35].

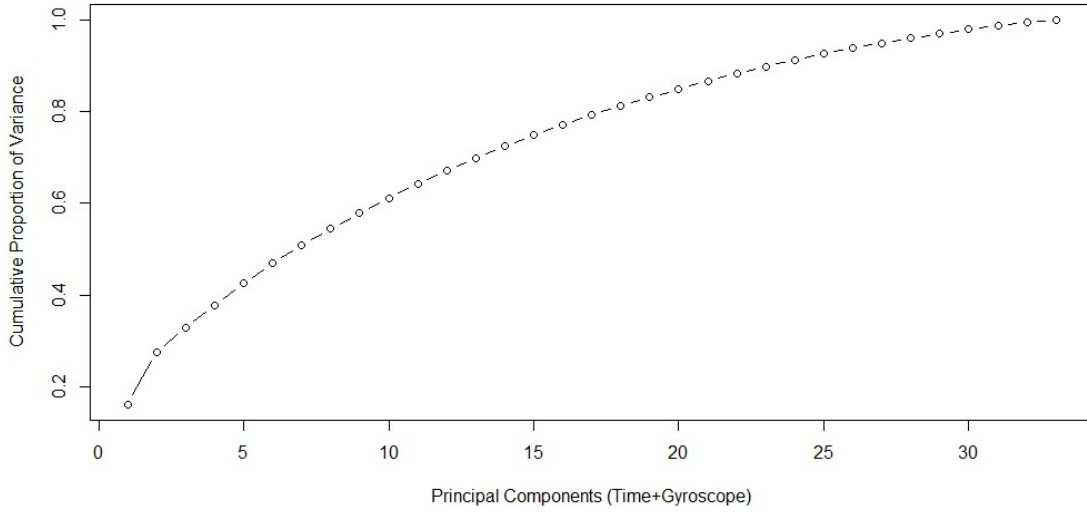
Temel Bileşen Analizi sonucunda farklı kombinasyonlardaki veri setlerinin özneliklerinin toplam varyans dağılım grafikleri aşağıdadır.



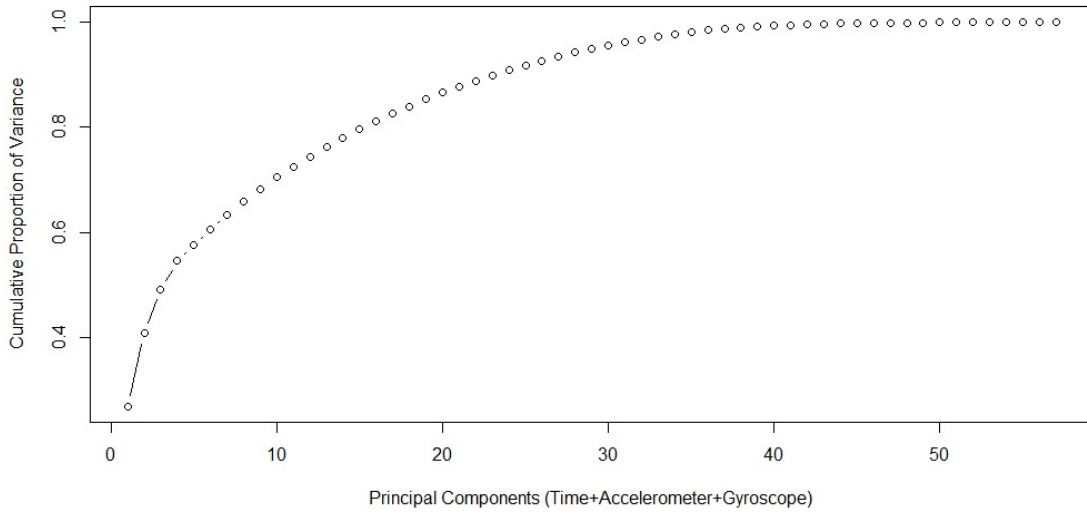
Şekil 22 Zaman Veri Seti Temel Bileşen Sayısı ve Kümülatif Toplam Varyans Grafiği



Şekil 23 Zaman ve Akselerometre Veri Seti Temel Bileşen Sayısı ve Kümülatif Toplam Varyans Grafiği



Şekil 24 Zaman ve Jiroskop Veri Seti Temel Bileşen Sayısı ve Kümülatif Varyans Grafiği



Şekil 25 Zaman, Akselerometre ve Jiroskop Veri Seti Temel Bileşen Sayısı ve Kümülatif Varyans Grafiği

Analizler esnasında kullanılacak temel bileşen sayısı belirtilmemiş olup otomatik olarak belirlenmesi sağlanmıştır.

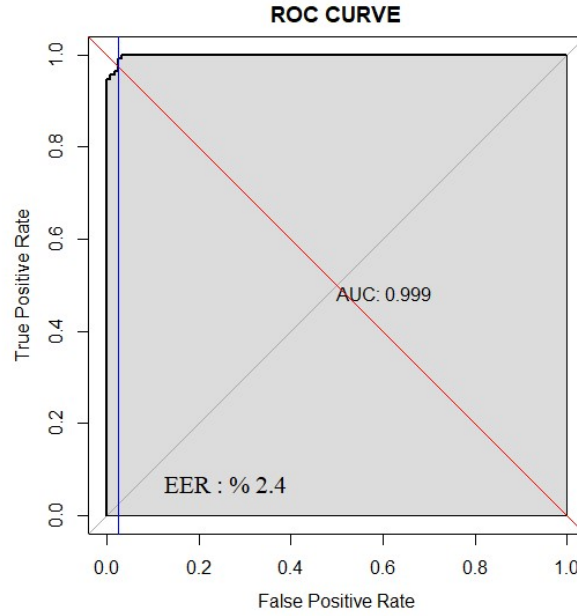
3.2.5.1. Zaman Veri Seti

YSA – Temel Bileşen Analizi yöntemiyle Zaman veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 25 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman Veri Setinin Sonuçları

Algoritma	Neural Network - PCA
FAR	1.60%
FRR	4.42%
EER	2.40%
AUC	0.999
Doğruluk	97.06%
Veri Seti	Zaman

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 26 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman Veri Setinin ROC Eğrisi ve EER Noktası

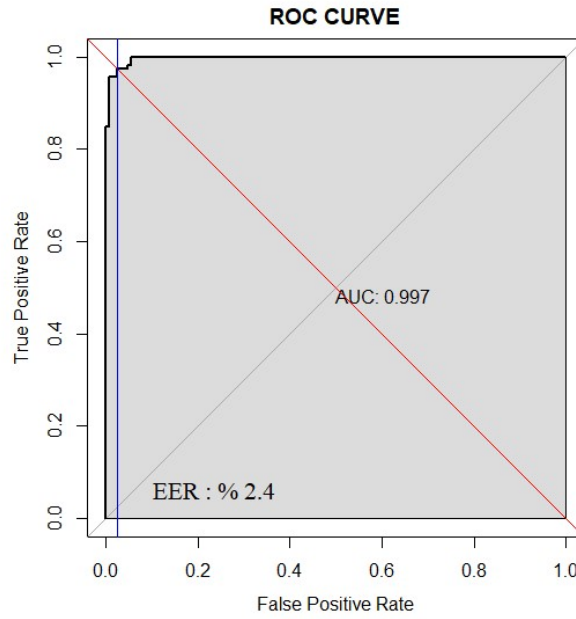
3.2.5.2. Zaman ve Akselerometre Veri Seti

YSA – Temel Bileşen Analizi yöntemiyle Zaman ve Akselerometre veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 26 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Akselerometre Veri Setinin Sonuçları

Algoritma	Neural Network - PCA
FAR	4.80%
FRR	2.65%
EER	2.40%
AUC	0.997
Doğruluk	96.22%
Veri Seti	Zaman ve Akselerometre

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 27 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Akselerometre Veri Setinin ROC Eğrisi ve EER Noktası

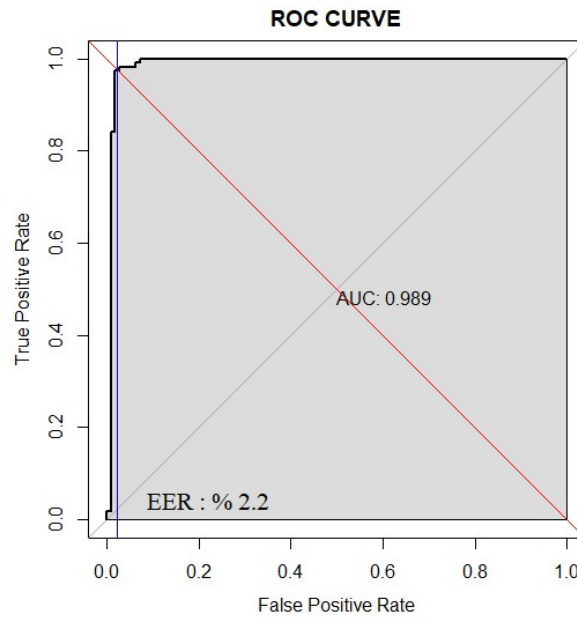
3.2.5.3. Zaman ve Jiroskop Veri Seti

YSA – Temel Bileşen Analizi yöntemiyle Zaman ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 27 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Jiroskop Veri Setinin Sonuçları

Algoritma	Neural Network - PCA
FAR	2.70%
FRR	1.77%
EER	2.20%
AUC	0.989
Doğruluk	97.77%
Veri Seti	Zaman ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 28 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

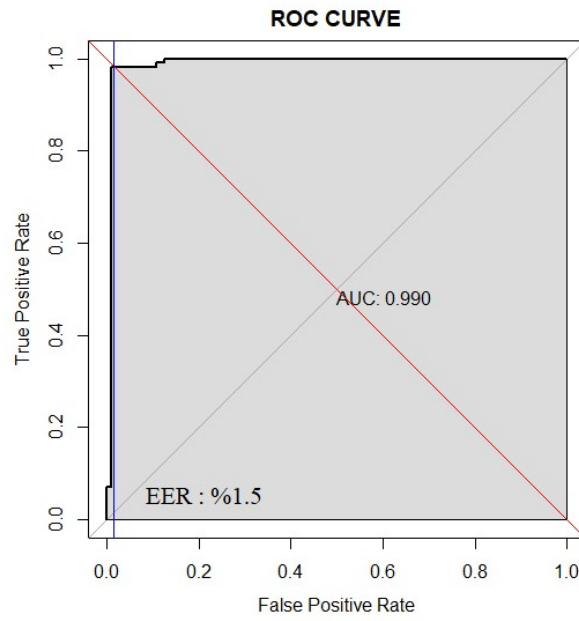
3.2.5.4. Zaman, Akselerometre ve Jiroskop Veri Seti

YSA – Temel Bileşen Analizi yöntemiyle Zaman ve Jiroskop veri seti 10 kez çapraz doğrulama yapılarak incelenmiştir. Elde edilen değerler aşağıdaki gibidir.

Tablo 28 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin Sonuçları

Algoritma	Neural Network - PCA
FAR	0.90%
FRR	1.77%
EER	1.50%
AUC	0.99
Doğruluk	98.66%
Veri Seti	Zaman, Akselerometre ve Jiroskop

ROC eğrisi ve EER noktası aşağıdaki gibidir.



Şekil 29 Yapay Sinir Ağları - Temel Bileşen Analizi Algoritmasıyla Zaman, Akselerometre ve Jiroskop Veri Setinin ROC Eğrisi ve EER Noktası

3.3. Ortalama Sonuçların Özeti

Bütün analizlerden (bkz: Bölüm 3.1) elde edilen değerler aşağıdaki tabloda toplanmıştır. Bu tablonun yorumlanmasına sonuçlarda (bkz: Bölüm 4) değinilecektir.

Tablo 29 Ortalama Sonuçların Özeti

Algoritma	FAR	FRR	EER	AUC	Doğruluk	Veri Seti
Karar Ağacı	11.20%	11.50%	11.50%	0.898	88.66%	Zaman
	8.00%	14.16%	11.20%	0.907	89.08%	Zaman+Aks.
	10.81%	7.96%	10.20%	0.921	90.62%	Zaman+Jir.
	12.07%	8.85%	11.20%	0.924	89.29%	Zaman+Aks.+Jir.
Naive Bayes	1.60%	8.85%	4.20%	0.976	94.96%	Zaman
	5.60%	6.59%	5.90%	0.991	91.20%	Zaman+Aks.
	2.70%	4.42%	2.60%	0.988	96.43%	Zaman+Jir.
	0.00%	6.19%	3.30%	0.996	96.88%	Zaman+Aks.+Jir.
Rastgele Orman	1.60%	2.65%	2.10%	0.997	97.90%	Zaman
	1.60%	4.42%	1.70%	0.999	97.06%	Zaman+Aks.
	0.90%	2.65%	1.90%	1	98.21%	Zaman+Jir.
	1.80%	0.88%	1.30%	1	98.66%	Zaman+Aks.+Jir.
YSA	1.60%	11.50%	7.50%	0.985	93.70%	Zaman
	4.00%	10.62%	8.40%	0.968	92.86%	Zaman+Aks.
	3.60%	8.85%	7.60%	0.95	93.75%	Zaman+Jir.
	9.01%	7.08%	8%	0.941	91.96%	Zaman+Aks.+Jir.
YSA - PCA	1.60%	4.42%	2.40%	0.999	97.06%	Zaman
	4.80%	2.65%	2.40%	0.997	96.22%	Zaman+Aks.
	2.70%	1.77%	2.20%	0.989	97.77%	Zaman+Jir.
	0.90%	1.77%	1.50%	0.99	98.66%	Zaman+Aks.+Jir.

3.4. Örneklem Sayına Göre Sonuç Karşılaştırma Çalışması

Bu çalışmada, örneklem sayısının sonuca olan etkisinin gözlemlenebilmesi için en iyi sonuç üreten Karar Ağacı Sınıflandırıcı ve en kötü sonuç üreten Rastgele Orman algoritmalarıyla Zaman, Akselerometre ve Jiroskop veri seti kullanılarak (bkz: Bölüm 4.1) her bir algoritma için 5 kez analiz yapıp ortalama doğruluk değerleri kaydedilmiştir. Veri setleri pozitif ve negatif etiket oranları korunarak (bkz: Bölüm 2.2.3), rastgele seçim yöntemiyle, toplam örneklem sayısının %25'i, %50'si, %75'i ve %100'ü olmak üzere 4 farklı grupta oluşturulmuştur.

Aşağıdaki tablonun yorumlanmasına sonuçlarda (bkz: Bölüm 4.1.4) değinilecektir.

Tablo 30 Örneklem Sayısına Göre Sonuçlar

Veri Seti Büyükülüğü	Algoritma	Doğruluk Oranı	FAR	FRR	Ortalama Doğruluk Oranı
25.00%	Rastgele Orman	100.00%	0.00%	0.00%	98.18%
25.00%	Rastgele Orman	98.18%	0.00%	3.57%	
25.00%	Rastgele Orman	96.36%	3.70%	3.57%	
25.00%	Rastgele Orman	100.00%	0.00%	0.00%	
25.00%	Rastgele Orman	96.36%	0.00%	7.27%	
50.00%	Rastgele Orman	99.10%	0.00%	1.79%	98.20%
50.00%	Rastgele Orman	97.30%	5.45%	0.00%	
50.00%	Rastgele Orman	99.10%	0.00%	1.79%	
50.00%	Rastgele Orman	98.20%	0.00%	3.57%	
50.00%	Rastgele Orman	97.30%	5.45%	0.00%	
75.00%	Rastgele Orman	98.80%	2.41%	0.00%	98.44%
75.00%	Rastgele Orman	96.41%	2.41%	4.76%	
75.00%	Rastgele Orman	99.40%	1.20%	0.00%	
75.00%	Rastgele Orman	99.40%	1.20%	0.00%	
75.00%	Rastgele Orman	98.20%	2.41%	1.19%	
100.00%	Rastgele Orman	97.32%	0.90%	4.42%	97.86%
100.00%	Rastgele Orman	97.77%	2.70%	1.77%	
100.00%	Rastgele Orman	97.32%	3.60%	1.77%	
100.00%	Rastgele Orman	98.66%	1.80%	0.88%	
100.00%	Rastgele Orman	98.21%	2.70%	0.88%	
25.00%	Karar Ağacı	87.27%	11.11%	14.29%	88.00%
25.00%	Karar Ağacı	85.45%	18.52%	10.71%	
25.00%	Karar Ağacı	89.09%	11.11%	10.71%	
25.00%	Karar Ağacı	90.91%	7.41%	10.71%	
25.00%	Karar Ağacı	87.27%	7.41%	17.86%	
50.00%	Karar Ağacı	90.99%	3.64%	14.29%	89.19%
50.00%	Karar Ağacı	92.79%	5.45%	8.93%	
50.00%	Karar Ağacı	84.68%	7.27%	23.21%	
50.00%	Karar Ağacı	87.39%	10.91%	14.29%	
50.00%	Karar Ağacı	90.09%	9.09%	10.71%	
75.00%	Karar Ağacı	88.62%	16.87%	5.95%	89.10%
75.00%	Karar Ağacı	88.02%	10.84%	13.10%	
75.00%	Karar Ağacı	89.22%	4.82%	16.67%	
75.00%	Karar Ağacı	89.22%	3.61%	17.86%	
75.00%	Karar Ağacı	90.42%	9.64%	9.52%	
100.00%	Karar Ağacı	89.29%	12.07%	8.85%	90.00%
100.00%	Karar Ağacı	91.07%	11.71%	6.19%	
100.00%	Karar Ağacı	91.96%	5.41%	10.62%	
100.00%	Karar Ağacı	87.50%	7.21%	17.70%	
100.00%	Karar Ağacı	90.18%	11.71%	7.96%	

3.5. Çoklu Sınıflandırma Deneysel Çalışması

Bu deneysel çalışmada, sistemin iki farklı kişiyi diğer kişilerden ayırt etme senaryosu en iyi sonuç üreten Karar Ağacı Sınıflandırıcı ve en kötü sonuç üreten Rastgele Orman algoritmalarıyla (bkz: Bölüm 4.1) Zaman, Akselerometre ve Jiroskop veri seti kullanılarak incelenmiştir.

Aşağıdaki tablonun yorumlanmasına sonuçlarda (bkz: Bölüm 4.1.5) değinilecektir.

Tablo 31 Çoklu Sınıflandırma Deneysel Çalışmasının Sonuçları

Algoritma	Doğruluk Oranı	FAR 1. kişi	FAR 2. kişi	FRR 1. kişi	FRR 2. kişi
Karar Ağacı	85.20%	12.39%	15.79%	6.36%	3.01%
Rastgele Orman	98.21%	0.00%	1.75%	3.64%	0.00%

4. SONUÇLAR

4.1. Sonuçların Değerlendirilmesi

Gerçekleştirilen bu tez çalışmasında, bir çeşit biyometrik kimlik doğrulama yöntemi olarak; tuşa basış dinamikleri ve mobil cihazlarda bulunan akselerometre ve jiroskop sensörlerinden toplanan verileri farklı makine öğrenmesi algoritmaları ile analiz edilmiştir. Bu bağlamda çalışmada elde edilen sonuçlar, farklı öz niteliklerin bulunduğu veri setlerinin sonuçlara etkisi ve kullanılan farklı makine öğrenmesi algoritmalarının sonuçlara etkisine göre iki farklı başlıkta EER metriği göz önünde bulundurularak değerlendirilmiştir. Ayrıca aşağıdaki başlıkta (bkz: Bölüm 4.1.3) tespit edilen en iyi sonuçlar belirtilmiştir. Ek olarak, Ortalama Sonuçların Özeti (bkz: Tablo 29) genel anlamda incelendiğinde; AUC ve Doğruluk oranı metrikleri EER baz alınarak yapılan kıyaslamaları destekler niteliktedir.

4.1.1. Farklı Veri Seti Kombinasyonlarının Sonuca Etkisi

Ortalama Sonuçların Özeti (bkz: Tablo 29) verilen ortalama EER değerleri baz alınarak yapılan değerlendirme sonucunda, farklı sensörlerden elde edilen farklı özelliklerdeki öz niteliklerin sonuca olan etkileri karşılaştırılmıştır.

Aşağıdaki tabloda veri setleri ve algoritmalara göre ortalama EER değerleri yer almaktadır. Öz nitelik bakımından veri seti zenginleştiği zaman performansı artan algoritmalar ve sonuçları aşağıda belirtilmiştir.

Tablo 32 Öz niteliklerin Sonuca Olan Etkileri

Veri Seti	Karar Ağacı	Rastgele Orman	Naïve Bayes	YSA	YSA - PCA
Zaman	11.50%	2.10%	4.20%	7.50%	2.40%
Zaman + Aks.	11.20%	1.70%	5.90%	8.40%	2.40%
Zaman + Jir.	10.20%	1.90%	2.60%	7.60%	2.20%
Zaman + Aks. + Jir.	11.20%	1.30%	3.30%	8%	1.50%

Veri seti öz nitelik sayısı bakımından zenginleşse de sonuçlarda dikkate alınabilecek bir iyileşme göstermeyen algoritmalar aşağıda belirtilmiştir;

- Karar Ağacı
- YSA

Veri seti öznitelik sayısı bakımından zenginleştikçe sonuçlarda iyileşme gözlemlenen algoritmalar aşağıda belirtilmiştir;

- Naive Bayes
- Rastgele Orman
- YSA - PCA

Öznitelik sayısı arttıkça sonuçlarda iyileşme gözlemlenen algoritmalar detaylı olarak incelendiğinde; bu iyileşmenin ‘Zaman’ veri setine eklenen her bir farklı sensöre ait öznitelik için farklı oranda gerçekleştiği, ancak özellikle ‘Zaman, Akselerometre ve Jiroskop’ verileri, yani farklı disiplinlere ait tüm veriler birlikte kullanıldığında aşağıdaki algoritmalarda daha belirgin olduğu gözlemlenmiştir;

- Rastgele Orman
- YSA - PCA

4.1.2. Farklı Makine Öğrenmesi Algoritmalarının Sonuca Etkisi

Veri setlerinin sonuca etkisi için yapılan değerlendirmelere (bkz: Bölüm 4.1.1.) ek olarak Ortalama Sonuçların Özetinde (bkz: Tablo 29) verilen ortalama EER değerleri baz alınarak yapılan değerlendirme sonucunda, farklı algoritmalar ile elde edilen sonuçlar karşılaştırılmıştır.

Aşağıdaki tabloda en iyi sonuçları üreten Zaman, Akselerometre ve Jiroskop veri setinin ortalama EER kullanılarak kıyaslama yapılmıştır. En iyi algoritma belirtilmiştir.

Tablo 33 Sınıflandırma Algoritmalarının Performansları

Algoritmalar	Düşük Performans	Orta Performans	Yüksek Performans
Karar Ağacı	11.20%	-	-
Rastgele Orman	-	-	1.30%
Naïve Bayes	-	3.30%	-
YSA	8%	-	-
YSA - PCA	-	-	1.50%

Analiz edilen tüm farklı algoritmalar arasından düşük performans gösterenler aşağıda belirtilmiştir;

- Karar Ağacı

- YSA

Analiz edilen tüm farklı algoritmalar arasından ortalama performans gösteren aşağıda belirtilmiştir;

- Naive Bayes

Analiz edilen tüm farklı algoritmalar arasından yüksek performans gösteren aşağıda belirtilmiştir;

- Rastgele Orman
- YSA - PCA

En iyi performans gösteren bu iki algoritma ile, veri setleri değerlendirmelerinde (bkz: Bölüm 4.1.1) belirtilen; öznelilik bakımından en zengin veri seti kullanılarak yapılan analiz sonucu en iyi değerleri üreten algoritmaların **aynı olduğu** görülmektedir.

4.1.3. En İyi Sonuçların Değerlendirilmesi

Veri setleri ve makine öğrenmesi algoritmalarının sonuca etkisinin değerlendirildiği bölümlerde (bkz: Bölüm 4.1.1 ve Bölüm 4.1.2) açıklanan en yüksek performans gösteren veri seti ve algoritma birlikte değerlendirildiğinde; ‘Rastgele Orman’ algoritması ve ‘Zaman, Akselerometre ve Jiroskop’ veri seti kullanılarak yapılan analiz ile en iyi ortalama sonuç olan %1.30 EER değeri tespit edilmiştir.

4.1.4. Örneklem Sayısına Göre Sonuçların Değerlendirilmesi

Tablo 30’daki Ortalama Doğruluk Değer’leri değerleri göz önünde bulundurulduğunda;

- Karar Ağacı Sınıflandırıcı için örneklem sayısı arttıkça Ortalama Doğruluk Oranında eser miktarda artışlar gözlemlenmiştir.
- Rastgele Orman Sınıflandırıcı için örneklem sayısına bağlı olarak kayda değer değişimler gözlemlenememiştir.

4.1.5. Çoklu Sınıflandırma Deneysel Çalışmasının Değerlendirilmesi

Tablo 29 ve Tablo 31’deki değerler göz önünde bulundurulduğunda, çoklu sınıflandırmanın vermiş olduğu sonuçlar, tasarlanan sistemin yüksek oranda doğru sonuç verdiğini destekler niteliktedir. Bu bağlamda kimlik doğrulama için geliştirilen bu

model, ihtiyaç duyulduğu takdirde, birden fazla kişinin ayırt edilmesi gereken durumlarda da kullanılabilir.

4.2. Siber Tehditler

Literatürde yer alan çalışmalar incelendiğinde, yalnızca sensör verileri kullanarak yüksek oranda kimlik doğrulama yapılabildiğini ortaya koyan çalışmalar mevcuttur [14-15, 17-18, 21-23] . Bu tez çalışmasında ise, hem sensör hem de tuşa basış dinamikleri kullanılmakta olup, tasarlanan sistemin atlatılabilmesi için hem sensör hem de tuşa basış dinamiklerinin ele geçirilmesine ihtiyaç duyulmaktadır.

Başka literatür çalışmaları incelendiğinde ise sensör verileri analiz edilerek hangi tuşa basıldığını tahmin eden çalışmaların da yer aldığı görülmüştür [24, 29-30]. Bu nedenle, dolaylı olarak da olsa basılan tuş bilgilerinin elde edilebileceği hesaba katılmalıdır. Bu çalışmalarda başarı yüksek orana ulaştığı zaman, bu tez çalışmasında tasarlanan davranışsal biyometrik kimlik doğrulama sistemi, hem sensör hem de sensörler sayesinde basılan tuşlar tespit edilebildiği için tehlikeye girmektedir.

Mobil cihazlarda bulunan sensörlerden veri toplanırken, kullanıcılardan sensörlere erişim izin istenmemektedir [24]. Günümüz dünyasında sensör verileri analiz edilerek, bu tez çalışması aracılığıyla dahi kimlik doğrulaması yüksek oranda başarılabilmiş olup; sensör verileri yorumlandığı zaman, kişiye özgü veriler elde edilebilmektedir. Mobil cihaz işletim sistemi üreticileri tarafından, kullanıcılara ait davranışsal biyometrik verilerin çalınmaması adına sensörlerden veri okunurken, kullanıcılardan uygulama bazlı izin istemenin zorunlu kılınması gerekmektedir. Bu geliştirme yapıldığı takdirde tasarlanan bu sistem güvenli olarak kullanılabilir.

Bunların dışında, mobil uygulamaya bulaşabilecek bir zararlı yazılım aracılığıyla sensör ve tuşa basış dinamikleri verilerinin toplanması ihtimali; tıpkı zararlı yazılımlarla şifrelerin çalınması senaryosu ile benzer olasılıktadır. Saldırganlar elde ettikleri şifreleri doğrudan kullanabilirken; burada elde edilen sensör ve tuşa basış dinamikleri verilerini davranışsal biyometrik kimlik doğrulama sistemini atlatabilecek şekilde yorumlayıp, sistemin beklediği formatta sunmaları gerekeceğinden, dolaylı olarak kullanabileceklerdir.

4.3. Kullanılabilecek Alanlar

Bu tez çalışmasında açıklanan kimlik doğrulama sistemi, verilerin elde edilebileceği her türlü mobil cihaz üzerinde ve kimlik doğrulama gerektiren her uygulama için kullanılabilir.

Bu sistem doğrudan kimlik doğrulama amaçlı kullanılabileceği gibi; kullanıcıya ek bir efor getirmeyeceğinden, mevcut kimlik doğrulama sistemleri için destekleyici faktör olarak da konumlandırılabilir.

Kimlik doğrulama dışında belli risk seviyesi üzerindeki işlemler; örneğin bankacılık uygulamaları aracılığıyla gerçekleştirilen finansal talimatlar için, bu talimatın risk seviyesine uygun yetkilendirme amacıyla kullanılabilir.

4.4. Gelecekteki Çalışmalar

Bu çalışmada toplanan veriler, sabit bir metnin yazımı aşamasında toplanmıştır. Kullanıcıların sabit bir metin ile sınırlandırılmadığı, özgürce istediklerini yazdığı ve daha uzun süre gözlemlenen veriler ile yeni analizler yapılabilir.

Mobil cihaz sektörünün gelişmesi ile birlikte, cihazlara eklenebilecek yeni donanımsal sensörlerin verileri de bu çalışmaya dahil edilebilir.

KAYNAKLAR

- [1] Gümüş F., Ata O., Balık H. H., (2018). Davranışsal Biyometrinin 5 Yılı: Kimlik Doğrulama ve Anomali Tespit Uygulamaları. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 30, 345-364.
- [2] Galbally J., McCool C., Fierrez J., Marcel S., Garcia J. O., (2009). On the vulnerability of face verification systems to hill-climbing attacks. Pattern Recognition, 43 (2010), 1027-1038.
- [3] Espinoza M., Champod C., Margot P., (2010). Vulnerabilities of fingerprint reader to fake fingerprints attacks. Forensic Science International, 204 (2011), 41-49.
- [4] Mahfouz A., Mahmoud T. M., Eldin A. S., (2017). A survey on behavioral biometric authentication on smartphones. Journal of Information Security and Applications, 37 (2017), 28-37.
- [5] Alotaibi S., Furnell S., Clarke N., (2015). Transparent Authentication Systems for Mobile Device Security: A Review. The 10th International Conference for Internet Technology and Secured Transactions (ICITST-2015), 14-15 December 2015, London, England.
- [6] Özen Z. (2016) Kimlik Doğrulaması İçin Tuş ve Vuruş Dinamiklerine Dayalı Bir Güvenlik Sisteminin Yapay Sinir Ağları İle Geliştirilmesi. Doktora Tezi, İstanbul Üniversitesi, İstanbul, Türkiye, 84-88,184.
- [7] Agun N. (2016) Biyometrik Kimlik Doğrulama İçin Alan Tabanlı Tuşa Basma Dinamikleri Analizi. Yüksek Lisans Tezi, Anadolu Üniversitesi, Eskişehir, Türkiye, 25-30, 47-54.
- [8] Uzun Y. (2013) User Authentication And Distinguishing Child Users From Adults With Keystroke Dynamics. PhD Thesis, Middle East Technical University, Ankara, Turkey, 21-31, 43, 64, 67.
- [9] Fridman L. (2014) Learning of Identity from Behavioral Biometrics for Active Authentication. PhD Thesis, Drexel University, Pennsylvania, USA, 8-22, 31-44.

- [10] Çeker H. (2017) Keystroke Dynamics for Enhanced User Recognition in Active Authentication. PhD Thesis, State University of New York at Buffalo, New York, USA, 21-44, 61-79, 95-96.
- [11] Moliono Y., Ham H., Darmawan D., (2018). Keystroke Dynamic Classification using Machine Learning Password Authorization. 3rd International Conference on Computer Science and Computational Intelligence 2018, 07-08 September 2018, Jakarta, Indonesia.
- [12] Krishnamoorthy S. (2018) Identification of User Behavioral Biometrics for Authentication using Keystroke Dynamics and Machine Learning. MSc Thesis, University of Windsor, Ontario, Canada, 23, 27, 64.
- [13] Syed Z.A. (2014) Keystroke and Touch-dynamics Based Authentication for Desktop and Mobile Devices. PhD Thesis. West Virginia University. Virginia, USA, 19-23, 23-47, 67.
- [14] Yoneda K. (2017) Mobile Sensor-Based Biometrics From Common Daily Activities. MSc Thesis. Fordham University, New York, USA, 12-14, 19.
- [15] Catal C., Tufekci S., Pirmit E., Kocabag G., (2015). On the use of ensemble of classifiers for accelerometer-based activity recognition. *Applied Soft Computing*, 37 (2015), 1018-1022.
- [16] Syed Z., Helmick J., Banerjee S., Cukic B., (2018). Touch gesture-based authentication on mobile devices: The effects of user posture, device size, configuration, and inter-session variability. *The Journal of System and Software*, 149 (2019), 158-173.
- [17] Maghsoudi J. (2018) A Behavioral Biometrics User Authentication Study Using Motion Data from Android Smartphones. PhD Thesis, Pace University, New York, USA, 33-34, 81-84, 122.
- [18] Bhattarai A. (2018) Increasing Accuracy of Hand-Motion Based Continuous Authentication Systems. MSc Thesis, Tennessee Technological University, Tennessee, USA, 27-28, 41-51.
- [19] Mendizabal-Vazquez I., Santos-Sierra D., Guerra-Casanova J., Sanchez-Avila C., (2014). Supervised classification methods applied to keystroke dynamics through

mobile devices. 2014 International Carnahan Conference on Security Technology (ICCST), 13-16 October 2014, Rome, Italy.

[20] Yuksel A. S., Senel F. A., Cankaya I. A., (2019). Classification of Soft Keyboard Typing Behaviors Using Mobile Device Sensors with Machine Learning. *Arabian Journal for Science and Engineering*, 44, 3929-3942.

[21] Karakaya N., Işıklar Alptekin G., Durmaz Incel O., (2019). Using behavioral biometric sensors of mobile phones for user authentication. 23rd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, 04-06 September 2019, Budapest, Hungary.

[22] Shen C., Chen Y., Guan X., (2017). Performance evaluation of implicit smartphones via sensor-based analysis. *Information Sciences*, 430-431 (2018), 538-553.

[23] Buriro A., Crispo B., Conti M., (2019). ANSWERAUTH: A bimodal behavioral biometric-based user authentication scheme for smartphones. *Journal of Information Security and Applications*, 44 (2019), 89-103.

[24] Lee W. (2019) User Authentication And Security Vulnerabilities Using Smartphone Sensors And Machine Learning, PhD Thesis, Princeton University, New Jersey, USA, 54-60, 71-73, 75-78, 80-108, 113-128.

[25] Singh M., Singh R., Ross A., (2019). A comprehensive overview of biometric fusion. *Information Fusion*, 52 (2019), 187-205.

[26] Lee H., Hwang J. Y., Kim D. I., Lee S., Lee S. H., Shin J. S., (2018). Understanding Keystroke Dynamics for Smartphones Users Authentication and Keystroke Dynamics on Smartphones Build-In Motion Sensors. *Security and Communication Networks*, Volume 2018, Article ID 2567463, 1-10.

[27] Coakley M. J. (2016) Keystroke Biometric Studies with Short Numeric Input on Smartphones. PhD Thesis, Pace University, New York, USA, 42, 82-91.

[28] Ahmadzaded E. (2018) Data Mining Algorithms for Decision Support Based on User Activities. PhD Thesis. Florida Institute of Technology, Florida, USA, 19, 124-133.

[29] Wang L. (2019) I Know What You Type on Your Phone: Keystroke Inference on Android Device Using Deep Learning, MSc Thesis, University of Kansas, Kansas, USA, 10, 23, 51-53.

[30] Yang Z., Zhao R., Yue C., (2018). Effective Mobile Web User Fingerprinting via Motion Sensors. 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering, 01-03 August 2018, New York, NY, USA.

[31] Android Sensors Overview :

(https://developer.android.com/guide/topics/sensors/sensors_overview) (Eriřim tarihi: 06.01.2020).

[32] Rhodes-Ousley M., (2013). The Complete Reference: Information Security, Second Edition, McGraw-Hill Education, New York, USA.

[33] About R :

(<https://www.r-project.org/about.html>) (Eriřim tarihi: 06.01.2020) .

[34] Lantz B., (2013). Machine Learning with R, Packt Publishing Ltd., Birmingham, United Kingdom.

[35] Caret Documentation :

(<https://cran.r-project.org/web/packages/caret/caret.pdf>) (Eriřim tarihi: 06.01.2020).

ÖZGEÇMİŞ

Umut Berhan BALKIR 1992 yılında İstanbul'da doğmuştur. Lise eğitimini 2006-2010 yılları arasında Hüseyin Yıldız Anadolu Lisesi'nde tamamlamıştır. 2015 yılında İstanbul Üniversitesi Bilgisayar Mühendisliği bölümünü bitirmiştir. Lisans öğreniminin ilk yıllarından itibaren Yazılım Mühendisliği ve Bilgi Güvenliği alanlarında çalışmıştır. Özel bir bankada çalışma hayatına devam etmektedir.

