



**MAKİNE ÖĞRENMESİ YÖNTEMLERİ KULLANARAK AKILLI
TELEFON VERİLERİ İLE GÜVENLİK PERSONELİNİN AKTİVİTE
TESPİTİ**

Serdar ASARKAYA

**YÜKSEK LİSANS TEZİ
SAVUNMA TEKNOLOJİLERİ ANA BİLİM DALI**

**SİVAS BİLİM VE TEKNOLOJİ ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

HAZİRAN 2022

ETİK BEYAN

Sivas Bilim ve Teknoloji Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

.....
Serdar ASARKAYA
24/06/2022

MAKİNE ÖĞRENMEŞİ YÖNTEMLERİ KULLANARAK AKILLI TELEFON VERİLERİ İLE GÜVENLİK PERSONELİNİN AKTİVİTE TESPİTİ

(Yüksek Lisans Tezi)

Serdar ASARKAYA

SİVAS BİLİM ve TEKNOLOJİ ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Haziran 2022

ÖZET

Ülkemizde güvenlik personelleri arazi, açık veya kapalı ortamlar gibi zorlu şartlarda görev yapmaktadırlar. Görev sırasında personelin güvenliğinin sağlanması en önemli esaslardan biridir. Bu kapsamda personelin görev sırasında farkı teknolojilerle izlenmesi söz konusudur. Bu çalışmada güvenlik personelinin görev sırasında aktivitelerinin izlenmesine olanak sağlayacak yardımcı teknolojiler incelenmiştir. Bunun için akıllı telefonlarda bulunan sensör verileri ile makine öğrenmesi yöntemleri kullanılarak aktivite tespiti olanakları araştırılmıştır. Araştırmada 19-48 yaş arası, akıllı telefon taşıyan 30 katılımcıdan elde edilen sensör verileri kullanılmıştır. Her katılımcıdan yürüme, tırmanma, inme, oturma, bekleme, uzanma olmak üzere 6 farklı aktivite elde edilmiştir. Veriler makine öğrenmesinde kullanılan, Destek Vektör Makinesi (SVM), Rastgele Orman (RF), En Yakın Komşuluk (KNN), Çok Katmanlı Algılayıcı (MLP) algoritmaları ile oluşturulan modellerde kullanılmıştır. Elde edilen sınıflandırma sonuçları hassaslık, özgüllük, doğruluk gibi metriklerle birlikte tablolar halinde verilmiştir. Veri tutarlılığını izleme açısından Alıcı Çalışma Karakteristiği Eğrisi (ROC Curve) her model için çizilmiştir. Sınıflandırma sonuçlarındaki farklılıkları incelemek amacıyla veri seti sırasıyla %90, %80, %70 eğitim ve %10, %20, %30 test verisi olmak üzere bölünerek modellerde değerlendirilmiştir. Eğitim ve test verileri için sonuçlar hata matrisleriyle ve tablolar halinde gösterilmiştir.

Bilim Kodu : 10425687

Anahtar Kelimeler : Makine Öğrenmesi, İot, Güvenlik Personeli, Sensör, İvme Ölçer, Jiroskop, Aktivite Tanıma, KNN, SVM, MLP, RF

Sayfa Adedi : 82

Danışman : Dr. Öğr. Üyesi Emre ÜNSAL

ACTIVITY RECOGNITION FOR SECURITY PERSONEL USING MACHINE
LEARNING MODELS WITH SMARTPHONE DATA

(M. Sc. Thesis)

Serdar ASARKAYA

SİVAS UNIVERSITY OF SCIENCE AND TECHNOLOGY
INSTITUTE OF GRADUATE STUDIES

June 2022

ABSTRACT

In our country, security personnel work in difficult conditions such as terrain, open or closed environments. Ensuring the safety of personnel while on duty is one of the most important principles. In this context, it is possible to monitor the personnel with different technologies while on duty. In this study, assistive technologies that will enable the monitoring of the activities of security personnel during the task are examined. For this, the possibilities of activity recognition were investigated by using sensor data on smart phones and machine learning methods. In the research, sensor data obtained from 30 participants aged between 19-48 and carrying a smart phone were used. 6 different activities were obtained from each participant: walking, climbing, descending, sitting, waiting, reaching. The data were used in models created with Support Vector Machine (SVM), Random Forest (RF), Nearest Neighborhood (KNN), Multilayer Perceptron (MLP) algorithms used in machine learning. Obtained classification results are given in tables together with metrics such as sensitivity, specificity and accuracy. The Receiver Operating Characteristic Curve (ROC Curve) was plotted for each model to monitor data consistency. In order to examine the differences in the classification results, the data set was divided into 90%, 80%, 70% training and 10%, 20%, 30% test data, respectively, and evaluated in the models. The results for the training and test data are shown with confusion matrices and in tabular form.

Science Code : 10425687

Key Words : Machine Learning, Iot, Security Personel, Activity, Sensor, Jyroscope, Accelerometer, Activity Recognition, KNN, SVM, MLP, RF

Page Number : 82

Supervisor : Dr. Öğr. Üyesi Emre ÜNSAL

TEŞEKKÜR

Çalışmamın araştırma ve uygulama aşamalarında fikirleriyle yön veren, tezime danışmanlık yapan çok değerli hocam Dr. Emre ÜNSAL'a ve manevi destekleriyle sürekli yanımda olan anneme ve babama, yoğun çalışma süreçlerinde anlayış gösteren biricik kızım Zeynep Sude ASARKAYA' ya şükranlarımı borç bilirim.



İÇİNDEKİLER

Sayfa

ÖZET	vii
ABSTRACT.....	viii
TEŞEKKÜR.....	ix
İÇİNDEKİLER	x
ÇİZELGELERİN LİSTESİ.....	xii
ŞEKİLLERİN LİSTESİ.....	xiii
SİMGELER VE KISALTMALAR.....	xiv
1. GİRİŞ.....	1
2. KAYNAK ARAŞTIRMASI.....	2
3. MATERYAL VE YÖNTEM.....	9
3.1 Veri Seti	9
3.2 Veri Ön İşleme	10
3.3 Jiroskop	11
3.4 İvme Ölçer.....	11
3.5 Analiz Yöntemleri	11
3.5.1 Makine öğrenmesi.....	12
3.5.2 Rastgele orman sınıflandırıcı (random forest)	13
3.5.3 Destek vektör makinesi (support vector machine).....	14
3.5.4 En yakın komşuluk (k-nearest neighbour)	15
3.5.4 Çok katmanlı algılayıcı (multi layer perceptron).....	16
3.5.5 Derin öğrenme (deep learning)	17
3.5.6 ROC eğrileri ve değerlendirme metrikleri	19
3.5.7 Çapraz doğrulama (cross validation)	21

Sayfa

3.5.8 Hata matrisi (confusion matrix)	21
4. BULGULAR VE TARTIŞMA	23
4.1. KNN Sınıflandırıcısı ile Elde Edilen Sonuçlar	23
4.1.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi	23
4.1.2 Test verisi ile metrikler	27
4.1.3 Test verisi ile roc eğrisi	29
4.2 SVM Sınıflandırıcısı ile Elde Edilen Sonuçlar	30
4.2.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi	30
4.2.2 Test verisi ile metrikler	33
4.2.3 Test verisi ile roc eğrileri	35
4.3 MLP Sınıflandırıcısı ile Elde Edilen Sonuçlar	37
4.3.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi	37
4.3.2 Test verisi ile metrikler	40
4.3.3 Test verisi ile roc eğrisi	42
4.4 RF Sınıflandırıcısı ile Elde Edilen Sonuçlar ve ROC Eğrisi.....	43
4.4.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi	44
4.4.2 Test verisi ile metrikler	47
4.4.3 Test verisi ile roc eğrileri	49
4.5 Derin Öğrenme Modeli ile Elde Edilen Sonuçlar ve ROC Eğrisi.....	50
4.5.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi	51
4.5.2 Test verisi ile metrikler	54
4.5.3 Test verisi ile roc eğrileri	54
5. SONUÇ VE ÖNERİLER.....	56
6.KAYNAKLAR	59
ÖZGEÇMİŞ	62

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3.2. Sensörlerden elde edilen veriler.....	9
Çizelge 3.3. Özellik çıkarmada kullanılan fonksiyonlar.....	10
Çizelge 4.1. KNN ile elde edilen hata matrisleri örneklem gösterimi	26
Çizelge 4.2 KNN modeli ile elde edilen metrikler	27
Çizelge 4.4. SVM modeli ile elde edilen metrikler	34
Çizelge 4.5. MLP ile elde edilen hata matrisleri örneklem gösterimi	39
Çizelge 4.6. MLP modeli ile elde edilen metrikler	41
Çizelge 4.7. RF ile elde edilen hata matrisleri örneklem gösterimi.....	46
Çizelge 4.8. RF modeli ile elde edilen metrikler	47
Çizelge 4.9. Derin öğrenme modeli ile elde edilen hata matrisleri örneklem gösterimi	53
Çizelge 4.10. Derin öğrenme modeli ile elde edilen metrikler	54
Çizelge 5.2. Aktiviteler ve algoritmaların %20 test verisi ile karşılaştırılması	57
Çizelge 5.3. Aktiviteler ve algoritmaların %30 test verisi ile karşılaştırılması	57

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 3.1. Rastgele orman sınıflandırıcı.....	14
Şekil 3.2. SVM parametreye bağlı düzlem oluşumu	15
Şekil 3.3. K-en yakın komşuluk sınıflandırması.....	16
Şekil 3.4. Doğrusal olmayan bir yapay sinir ağı modeli.....	17
Şekil 3.5. Derin öğrenme model mimarisi.....	18
Şekil 3.6. Roc eğrisi	19
Şekil 3.7. Precision ve recall metrikleri örnek gösterim.....	20
Şekil 3.8. Eğitim verileri üzerinde çapraz doğrulama.....	21
Şekil 3.9. Gerçek ve tahminleri içeren bir hata matrisi.....	22
Şekil 4.1. KNN algoritması hata matrisleri a) %10 test b) %20 test c) %30 test	24
Şekil 4.2. KNN modeli ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test	29
Şekil 4.3. SVM algoritması hata matrisleri a)%10 test b)%20 test c)%30 test	31
Şekil 4.4. SVM modeli ile elde edilen roc eğrileri a)%10 test b) %20 test c)%30 test	36
Şekil 4.5. MLP algoritması hata matrisleri a)%10 test b)%20 test c)%30 test.....	38
Şekil 4.6. MLP modeli ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test.....	43
Şekil 4.7. RF algoritması hata matrisleri a)%10 test b)%20 test c)%30 test	44
Şekil 4.8. RF modeli ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test.....	49
Şekil 4.9. Derin Öğrenme Modeli Eğitim ve Doğruluk Eğrileri.....	50
Şekil 4.10. Derin öğrenme modeli hata matrisleri a)%10 test b)%20 test c)%30 test.....	51
Şekil 4.11. Derin öğrenme ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test...	55

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

Σ

Açıklamalar

Toplam

x

Veri seti girdileri

y

Veri seti sınıf etiketleri

p

Göreceli sınıf frekansı

c

Veri setindeki sınıf sayısı

w

Nöron ağırlığı

b

Bias değeri

ϕ

Aktivasyon fonksiyonu

u

Doğrusal toplam

y

Nöron çıkış sinyali

Kısaltmalar

Açıklamalar

AUC

Roc eğrisi altında kalan alan

CFS

Korelasyon tabanlı özellik seçimi

CNN

Evrişimli yapay sinir ağı

CRF

Koşullu rastgele alanlar

DNN

Derin sinir ağları

FPR

Yanlış pozitif oranı

FSS

Özellik alt küme seçimi

KNN

En yakın komşuluk sınıflandırıcısı

LDA

Lineer diskriminant analizi

LSTM

Uzun kısa süreli hafıza

MLP

Çok katmanlı algılayıcı sınıflandırıcısı

PCA	Temel bileşen analizi
RBF	Radial tabanlı fonksiyon
RF	Rastgele orman sınıflandırıcısı
RFID	Radyo frekans tanımlama
RNN	Tekrarlayan sinir ağları
ROC	Alıcı çalışma karakteristiği
SVM	Destek vektör makinesi sınıflandırıcısı
TPR	Doğru pozitif oranı
YSA	Yapay sinir ağı



1. GİRİŞ

Ülkemizde güvenlik personelleri yurt içi ve yurt dışı görevlerini, meskûn mahal, kapalı ortam, arazi vb. farklı ortamlarda başarı ile icra etmektedirler. Personelin görev yaptığı sırada izlenmesi ve güvenliklerinin sağlanması ve sahadan anlık bilgi akışının takibi de önemli bir husustur. Bu amaçla GPS sinyalleri, insanlı-insansız hava araçları gibi yöntemler kullanılabilir. Bu yöntemler personellerin konumu ve görev durumu hakkında fikir verirken aktiviteleri hakkında detaylı bilgi sunmamaktadırlar.

Giyilebilir teknolojilerin kullanımı, güvenlik personellerinin sınır içi, sınır dışı ve meskûn mahal operasyonlarında büyük yararlar sağlayabilmektedir. Giyilebilir teknolojiler, üzerinde sensör bulunduran ve anlık veri iletimi sağlayan cihazlar olarak adlandırılabilir.

ASELSAN'ın üzerinde çalıştığı CENKER projesi, askeri yürüyüş asistanı vb. projeler muharebe alanında insanlı-insansız kara unsurları arasında iki yönlü gerçek zamanlı veri akışını sağlayarak komuta kontrol seviyesini en üst düzeye çıkarmayı amaçlamaktadır. Güvenlik personellerinin üzerinde taşıyacağı akıllı telefon, saat vb. sensör içeren cihazlara ait verilerin makine öğrenmesi yöntemleriyle analizinin yapılması ilgili teknolojilerin geliştirilmesine katkı sağlayacağı söylenebilir.

Bu kapsamda alternatif olarak araştırmanın, güvenlik personellerinin yanlarında taşıyacağı akıllı telefonlara veya giyilebilir cihazlara ait sensör verilerinin makine öğrenmesi algoritmalarıyla değerlendirilerek, personellerin takibinin yapılmasına ve saha izlemesine katkı sağlayarak karar verme süreçlerinde yardımcı olarak kullanılabileceği düşünülmektedir.

Bu tez çalışmasında; akıllı telefonlarda bulunan ivmeölçer ve jiroskop sensör verileri kullanarak makine öğrenmesi yöntemleri ile güvenlik personellerinin fiziksel aktivitelerinin tanımlanması ortaya konmuştur.

2. KAYNAK ARAŞTIRMASI

Balkır (2020) yaptığı çalışmada, ivme ölçer ve jiroskop sensörlerinden toplanan verileri farklı makine öğrenmesi algoritmaları ile analiz etmiş ve akıllı telefonlarda üç boyutlu uzayda tuşlama verileri ile kimlik doğrulama modeli geliştirmiştir. Çalışmada karar ağacı (Decision Tree), Yapay Sinir Ağları (YSA), Naive Bayes, Rasgele Orman (Random Forest-RF) ve YSA-PCA algoritmaları ile oluşturulan modeller kullanılmıştır. Karar ağacı ve YSA en düşük performansı gösterirken, en iyi sonucu veren modelin %1,30 EER değeri ile Rasgele Orman algoritması modelinin olduğu belirtilmiştir.

Hernandez, Suzuki ve Venture (2020), Amerikan işaret diline yönelik aktivite tanımlama üzerine yaptıkları çalışmada, farklı konvolüsyonel yapay sinir ağları ve derin öğrenme algoritmalarını karşılaştırmışlardır. LeapMotion (Marin, Dominio, & Zanuttigh, 2014) veri seti ile yaptıkları araştırmada, Amerikan işaret dili tanımlamada en iyi sonucu %91,1 doğruluk oranı ile DeepConvLSTM modelinden elde ettiklerini belirtmişlerdir. ConvNet modeli ile %89,3 oranında doğruluk elde ederken, tekrarlayan sinir ağı(Recurrent Neural Network-RNN) modelinde bu oran %87,2' de kalmıştır. Veri üzerinde yapılan manipülasyonun en iyi sonuçta %3,8'lik bir doğruluk artışına yol açtığı vurgulanmıştır.

Singh ve diğerleri (2017) çalışmalarında, Naive Bayes, HMM, Hidden Semi Markov Model (HSMM), Conditional Random Fields (CRF) gibi standart makine öğrenmesi modelleri kullanarak insan aktivitesi tanımlamada elde ettikleri sonuçları Long-Short Term Memory (LSTM) ağlarından RNN modelinden elde ettikleri sonuçlarla karşılaştırmışlardır. LSTM ağında optimizasyon için Adam algoritması kullanırken, öğrenme hızını 0,0004 olarak seçmişlerdir. Tensorflow kütüphanesi kullanarak Long-Short Term Memory (LSTM) ağını modele eklemişlerdir. Sonuç olarak işlenmemiş ham verilerle desteklenen derin öğrenme modellerinin performansta iyileşme ve daha yüksek doğruluk oranlarına ulaşmada öncülük ettiği söylenebilir.

Lara ve Labrador (2013) çalışmalarında, giyilebilir sensör teknolojilerine dayalı insan aktivitelerini tanımlamada kullanılan yaklaşımları incelemişlerdir. Yanıt zamanı, esneklik, öğrenme yaklaşımları, doğru tanımlama açısından, makine öğrenmesi ve özellik çıkarmayı da içeren 28 farklı sistemi karşılaştırarak, gelecekte yapılacak çalışmalara ışık tutmuşlardır. Farklı çalışmaları incelerken, her çalışmada farklı veri setlerinin kullanıldığını, bazı çalışmalarda tek birey üzerinden veri toplanırken diğer çalışmalarda birden çok kişiye ait verilerin kullanıldığını belirtmişlerdir. Aktivite tanımlamada, göz önünde bulundurulanan

aktiviteler farklılık göstermektedir. Örneğin bazı çalışmalarda yürüme, koşma gibi günlük rutin aktiviteler tanımlanırken diğer bir çalışmada rutin dışı aktiviteler tanımlanmıştır. Çalışmalar arasındaki farklılıklardan biri de farklı yaş grupları ile çalışılmış olmasıdır.

Ranasinghe, Al Machot ve Mayr (2016)' e göre, aktivite tanımlama sistemleri, makine öğrenmesi algoritmaları ile donanım mimarisindeki gelişmelerin etkilediği geniş bir araştırma ve geliştirme alanını ifade etmektedir. Aktivite tanımlama çalışmaları, akıllı ev sistemleri, sağlık görüntüleme, iç ve dış mekân aktivitelerinin takibi amacıyla kullanılmaktadır. Bu uygulamalar farklı metotlara dayanmaktadır. Aktivite tanımlamada sanal yöntemlerle birlikte, sanal olmayan yöntemler ve çok modlu sensör teknolojileri kullanılmaktadır. Tüm bu uygulamaların farklı avantajları ve limitleri bulunmaktadır. Sanal sensör tabanlı çalışmalarda tek bir kamerayla geniş açılı bir ortam izleme avantajı varken, bu kameraların pahalı olması, yüksek maliyet ve yüksek işlem gücü gerektirmesi dezavantaj olarak karşımıza çıkmaktadır. Gerçek bir sensörle yapılan çalışmalarda insan vücuduna ait tüm aktiviteler gerçek zamanlı izlenebilir. Aynı zamanda düşük güç tüketimi, düşük maliyet ve hızlı veri işleme bu sistemlerin avantajı olarak ön plana çıkmaktadır. Tek bir sensörle yapılacak izlemeler hatalı tanımlamalara neden olabilir ve bu durum çalışmanın amacından sapmasına neden olabilir. Çok modlu sensörlerle yapılan çalışmalarda ise kompleks aktiviteler düşük güç tüketimi ve düşük maliyetlerle yapılabilirken, çoğu zaman hedef kişi üzerinde sürekli olarak birden çok sensör içeren akıllı kit taşımak zorundadır.

İnsan aktivitelerini tanımlamaya yönelik mobil/giyilebilir sensörler, ulaşılabilir olduklarından ve gizliliği ihmal etmeyen veriler ürettikleri için uzun yıllardır incelenmektedir. Giyilebilir teknolojilerle aktivite tanımlamada kullanılan teknolojilerden biri de, küçük boyutlu ve düşük maliyetli olmasının sağlamış olduğu avantajlardan dolayı Radyo Frekans Tanımlama (RFID) sistemleridir. Li, Marsic, Zhang ve Yang (2016) çalışmalarında, RFID kullanarak medikal ortamda bir travma odasında 7 farklı tıbbi aktiviteyi etiketleyerek veri toplamışlardır. 25 farklı gerçek travma vakasına müdahale sırasında toplanan veriler bir derin öğrenme modeline (DNN) sunularak aktivitelerin tespitinde ortalama %85 oranında başarı sağlanmıştır. Derin öğrenme modeline eklenen pooling katmanıyla başarı oranının %5 arttığı gözlenmiştir.

Rahimi ve diğerleri (2008) çalışmalarında, aktivite tanımlamada kullanılmak üzere, içerisinde 7 farklı sensör bulunduran Mobile Sensing Platform (MSP) adından bir sistem geliştirmişlerdir. MSP akış diyagramı içerisinde yürüme, koşma, tırmanma vb. aktivite sınıflandırmanın yanı sıra, özellik işleme adımlarına da yer verilmiştir. Çalışmada 12 farklı

katılımcıdan 207 fiziksel aktivite olayına ait veriler toplanmıştır. Zamansal bilgiden bağımsız toplanan ve tamamı etiketlenmiş verilerle eğitilen denetimli öğrenme modelinde %83,6 doğruluk oranı elde etmişlerdir. Aynı çalışmada zaman bilgisi ile veriler değerlendirildiğinde başarı oranı %93,8' e ulaşmıştır. Yarı denetimli modelde eğitim verilerinin %5' ile %79,7, %20' si ile %83,1, %40' ı ile %87,4 başarı oranları yakalanmıştır.

Berchtold, Budde, Gordon, Schmidtke ve Beigl (2010) yapmış oldukları çalışmada, akıllı telefonların giyilebilir olması nedeniyle aktivite tanımlamada kullanılabilir güçlü platformlar olduğunu ifade etmiştir. Akıllı telefonların yerleşik sensörlere sahip olmaları ve güçlü işlem birimleri sunmaları nedeniyle, sonraki adımda günlük etkinliklerin takip edilmesinde bu platformlara zekâ da eklenecektir. Yaptıkları çalışmada ActiServ adını verdikleri ve akıllı telefonlara aktivite tanımlamada yardımcı olacak bir servis yazmışlardır. ActiServ bileşenlerini telefonuna indiren kullanıcıdan aktivite verileri toplanarak ActiServ sunucusuna iletilmektedir. Sensörlerden toplanan veriler özellik çıkarım yöntemleriyle değerlendirildikten sonra aktivite sınıflandırma modülüne iletilir. Sınıflandırma modülleri M1, M2, M3, M4 olarak ayrıştırılmıştır. M1 modülü kullanıcının cebinde bulunan ama hareket etmediği durumdaki aktiviteleri tanımlamakta, M2 modülü kullanıcının telefonu cebinde taşıdığı ve hareket ettiği anda aktivite tanımlamakta, M3 modülü telefonun bir masa üzerinde hareket etmediği anda sensör verileri ile tanımlanmakta, M4 modülü ise kullanıcının telefonu elinde tuttuğu andaki aktiviteleri tanımlamaktadır. Belirtilen modülleri kullanarak yapılan sınıflandırmada %60-%97,3 arasında doğruluk oranlarına ulaşılmıştır.

İnsan jestlerini ve aktivitelerini giyilebilir sistemlerle tanıma giderek yaygınlaşmaktadır. Günümüzde birçok araştırmacı biyomedikal mühendislik, tıbbi hemşirelik, etkileşimli çözümler gibi alanlarda insan etkinliğini ve aktivitelerini tanıma teknolojileri üzerine odaklanmıştır. Aktivite tanıma amaçlı veriler elektromiyografi (EMG), ses sensörleri, görüntü sensörleri, akselerometre gibi kaynaklardan toplanabilmektedir. Akselerometre sensöründen toplanan verilerle Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Decision Tree, Naive Bayes, Hidden Markov Model gibi sınıflandırma algoritmaları kullanılarak aktivitelerin sınıflandırılması gerçekleştirilmektedir. Üç açılı ivmeölçer (akselerometre)'den toplanan verilerle yapılan çalışmada (Yang, Chen, Lee, Liou, & Wang, 2007), Feature Subset Selection (FSS) ve Linear Discriminant Analysis (LDA) yöntemleriyle özellik seçimi yapılmıştır. Veriler 5 katmanlı bir sınıflayıcıya sunularak, yürüme, koşma, bekleme, masaüstü bilgisayar ile çalışma, oturma, dış

fırçalama, süpürme gibi aktivitelerin sınıflandırma işlemi gerçekleştirilmiştir. İki farklı özellik seçim yöntemiyle %83 ve %93 oranlarında başarı elde edilmiştir.

Giyilebilir çoklu sensörlerinin vücuttaki pozisyonlarının aktivite tanımayaya etkisinin araştırıldığı çalışmada (Maurer, Smailagic, Siewiorek, & Deisher, 2006), verileri eWatch cihazı üzerinde bulunan ivmeölçer ve ışık sensörlerinden toplamışlardır. Zaman ve enerji tüketimini azaltmak amacıyla WEKA aracından özellik seçim yöntemlerinden Correlation Based Feature Selection (CFS) yöntemi tercih edilen çalışmada, karar ağacı, en yakın komşuluk ve Naive Bayes sınıflandırıcıları kullanılmıştır. Eğitim sırasında çapraz veri seçimi yapılarak doğrulama işlemi gerçekleştirilmiştir. Her bir vücut pozisyonu için sınıflandırma verileri değerlendirildiğinde, sensörler çantada taşındığında %92,8, cepte taşındığında %85,2, kimerde taşındığında %87, boyunda taşındığında %86,8 doğruluk oranlarına ulaşılmıştır. Yapılan çalışma Defense Advanced Research Projects Agency (DARPA) tarafından desteklenen çalışmalara dayanmaktadır.

Bao ve Intille (2004)' e göre aktivite tanımadada eğitim ve test verilerinin laboratuvar ortamı yerine doğal koşullarda toplanmış olması önem taşımaktadır. Laboratuvar ortamında toplanan veriler aktivite tanıma işlemini kısıtlayabilir, basitleştirebilir veya gerçek hayatta uyuşmayan sonuçlara neden olabilir. Araştırmacıların vücudun farklı bölgelerine yerleştirdikleri ivmeölçer, video veya ses sensörlerinin ürettiği verilerle aktivite tanıma amaçlı giyilebilir sistemler geliştirilmeye devam etmektedir. Söz konusu çalışmada ivmeölçer sensöründen toplanan verilerin değerlendirildiği modellerde %41,42 ile %97,49 arasında değişen doğruluk oranlarına ulaşılmıştır. Elde edilen değerler tv izleme, yürüme, bisiklet sürme, kitap okuma, merdiven çıkma, koşma, yeme, içme vb. farklı aktiviteleri içermektedir.

Ölçeklenebilir ve uyarlanabilir mobil sistemlerin ürettiği mobil big data, milyonlarca modelleme parametresi içeren derin öğrenme modelleri kullanarak analiz edilebilmektedir. Analiz edilen verilerden sonuç elde edebilmek için güçlü kaynaklara ve uzun zamana ihtiyaç duyulmaktadır.

Kaynak optimizasyonu ve zamanın verimli kullanımı amacıyla mobil big datanın analizini sağlayan bir çerçeve önerilmiştir. Söz konusu çerçeve açık kaynak hesaplama platformu olan Apache Spark üzerine inşa edilmiştir. Böylelikle sürekli erişilen verilerin ön belleğe alındığı bir veri kümesinde, birçok bilgi işlem çekirdeğinin kullanıldığı dağıtılmış öğrenme mümkün kılınabilir. Bu durum derin öğrenme modellerinin öğrenmesini birkaç kat hızlandıracaktır. Aktivite tanımadada derin öğrenme modellerinin de kullanılabileceği

düşünüldüğünde mobil cihaz verileri için hızlı ve yüksek doğruluk oranlarına ulaşmak mümkün olacaktır. (Alsheikh, Niyato, Lin, Tan, & Han, 2016)

Weiss ve Lockhart (2012)' e göre, akıllı telefonlardaki ivmeölçer sensörü aktivite tanıma çalışmalarında yaygın olarak kullanılmaktadır. Sensörden toplanan veriler kişisel veya kişisel olmayan şeklinde ayrıldığında aktivite tanıma çalışmalarına etkisi araştırılmış ve kişisel verilerle yapılan yürüme, tırmanma, uzanma, bekleme gibi aktivitelerin tanınmasında daha yüksek başarı oranlarına ulaşılmıştır.

Kişisel verilerle %98' den fazla başarı oranı yakalanırken, kişisel olmayan yani birden çok katılımcının verileri birlikte değerlendirildiğinde başarı oranları %68-%75,9 arasında kalmıştır. Çalışmada android işletim sistemine sahip akıllı telefon taşıyan 59 katılımcıdan toplanan verilerden yararlanılmıştır. Standart sınıflandırma algoritmalarının varsayılan parametrelerle kullanıldığı çalışmada %99,8 ile yapay sinir ağı ile koşu aktivitesinde en yüksek doğruluk oranına ulaşılmıştır. Verilerin karıştırılarak kullanıldığı sınıflandırmada en düşük başarı oranı %34 ile IB3 algoritması ile uzanma aktivitesinde elde edilmiştir. (Weiss ve diğerleri, 2012)

Györbiro, Fabian, & Homanyi (2009)' e göre kişisel iletişim araçları olan akıllı telefonlar, düşük enerji tüketimi ve üzerlerinde bulundukları sensörler sayesinde ürettikleri verilerle aktivite tanımada kullanılabilir. İvmeölçer sensörleri ise hareket algılamada kullanılabilen güçlü araçlardır. Yaptıkları çalışmada derin öğrenme ile yürüme, dinleme, yazma bisiklet sürme gibi aktivitelerin sınıflandırmışlardır. Veri seti üzerinde boyut azaltma yöntemi uygulayarak elde ettikleri sonuçlara göre dinlenme ve ellerin hareket ettiği durumlarda modelin başarı oranı düşerken, yazma, yürüme, koşma, bisiklet sürme aktivitelerinde başarı oranlarının arttığı gözlemlenmiştir.

Sakacı (2018) çalışmasında, 9 askeri personel üzerine yerleştirilen donanımlarla, arayüz yazılımı kullanarak aralarında sesli, yazılı ve görsel iletişim sağlayan askeri personel için kıyafet tasarımı gerçekleştirmiştir. Bu giyilebilir kıyafet ile askeri personelin konum bilgisinden görev bilgisine, ilk yardım gereksiniminden, kamera ile görsel iletişime kadar bir konuda askeri personelin izlenmesi gerçekleştirilmiştir. Modüler olarak geliştirilen sistem içerisinde GPS, kalp atışını algılayan sensör, kulaklık, mikrofon ve wi-fi cihazları bulunmaktadır. Günlük hayatta medikal amaçlı kullanılabilecek sistemin askeri personele uyarlanarak güvenlik güçlerinin takibi amacıyla da kullanılabileceği belirtilmiştir.

Giyilebilir uygulamalarda sensörlerden yararlanabilmek için 2 temel gereksinim karşılanmalıdır. Birincisi cihazın vücut üzerinde göze batmayacak ve hareket

engellemeyecek şekilde yerleştirilebilmesidir. İkinci olarak maliyeti düşük ve boyut olarak taşınabilir olması gerekmektedir.

Aktivite tanımada kaç adet sensörün kullanılması gerektiği ve vücudun hangi bölgesine yerleştirilmesi gerektiği soruları da cevaplanmalıdır. (Kern, Schiele, & Schmidt, 2003), yaptıkları çalışmada bir donanım platformu geliştirerek insan vücudundan eş zamanlı olarak ivmeölçer verisi toplamışlardır. Naive Bayes sınıflandırıcısı kullanarak ihtiyaç duyulan sensör sayısı ve sensörlerin yerleşimine göre aktivite tanıma çalışması gerçekleştirmişlerdir. Oturma, bekleme ve yürüme gibi aktivitelerin tanımlandığı çalışmada sensörler sırasıyla sol kol, sol bacak, sağ kol, sağ bacak üzerinde taşınarak elde edilen sonuçlar karşılaştırılmıştır. Daha kompleks aktivitelerin tanınmasında birden fazla sensör kullanımının daha iyi sonuçlar verdiği vurgulanmıştır.

Başlangıç ve varış noktalarını içeren yürüyüş aktivitesi tanımlanabildiğinde kişiye ait konum tahminleri de yapılabilmektedir. Kişinin attığı adım sayısı kullanarak elde edilen tahmini konum bilgileri, aktivite tanıma amaçlı geliştirilen sensör içeren donanımlarla mümkün olabilmektedir.

Lee ve Mase (2002)'e göre kapalı ortamda aktivite tabanlı konum belirleme çalışmalarında rota hataları konumun yanlış belirlenmesine neden olabilmektedir. Sorunun çözümü için 2B uzaydaki hareketin tanımına değil sözlü açıklamalara dayanan bir konum belirleme sistemi üzerinde çalışmışlardır. Adım sayıları takip edilerek konum tahmini yapılırken kişi varış noktasına ulaşmadan konum geçişleri ortaya çıkmıştır. Hata kişinin aktivitesinin yanlış tanınmasından kaynaklanmaktadır. Çalışmada bu sınırlılıkları ortadan kaldırmak amacıyla farklı sensörlerle farklı vücut pozisyonlarından elde ettikleri verileri kullanmışlardır.

Akıllı telefonlar ve akıllı saatlerin kullanımının yaygınlaşması, ivmeölçer ve jiroskop gibi sensörleri içerdikleri için günlük hayatta aktivite tanımayı kolaylaştırmaktadır. Klasik sınıflandırıcılar haricinde Evrişimli Yapay Sinir Ağları (Convolutional Neural Network-CNN) ve Uzun-Kısa Dönem Hafıza (LSTM) gibi derin öğrenme algoritmaları da aktivitelerin sınıflandırılmasında kullanılabilir.

Oluwalade, Neela, Wawira, Adejumo ve Purkayastha (2021), ivmeölçer ve jiroskop sensör verileri ile aktivite tanıma çalışmalarının verimli olduğunu belirtmişler ve kendi çalışmalarında Fordham Üniversitesi'nde Wireless Sensor Data Mining (WISDM) laboratuvarında ivmeölçer ve jiroskop sensörleri ile oluşturulan veri setini (Kwapisz, Weiss, & Moore, 2012) kullanmışlardır. Bilekte bulunan akıllı saat ve cepte taşınan akıllı telefon kullanılarak 3 dakikalık sürelerle 18 farklı aktiviteye ait veriler toplanarak WISDM

veriseti olarak adlandırılmıştır. Her bir sensör eksenine için standart sapma, ortalama, varyans gibi hesaplamalar yapılarak veri seti ön işleme işlemlerinden geçirilmiştir. Aktivite sınıflandırmada en yüksek başarı oranı %98 ile Evrişimli Sinir Ağı modeli ile elde edilmiştir.



3. MATERİYAL VE YÖNTEM

3.1 Veri Seti

Çalışmada, (Ortiz, Anguita, Ghio, Oneto, & Parra, 2021) tarafından insan aktiviteleri ve hareket bozukluklarını tanıma amacıyla yapılan çalışmada kullanılan, ivmeölçer ve jiroskop sensörleri ile üretilen akıllı telefon verileri kullanılmıştır. Veri seti, 19-48 yaş arası, akıllı telefon taşıyan 30 katılımcıdan elde edilen sensör verilerini içermektedir. Her katılımcıdan yürüme, tırmanma, inme, oturma, bekleme, uzanma olmak üzere 6 farklı aktivite elde edilmiştir. Her bir kayıt bu aktivitelerle etiketlenmiştir. Aktiviteler ve etiketler Çizelge 3.1’de gösterilmiştir.

Çizelge 3.1. Aktiviteler ve etiketler

Aktivite	Etiket
Yürüme	0
Tırmanma	1
İnme	2
Oturma	3
Bekleme	4
Uzanma	5

Veri setinde, ivmeölçer ve jiroskop sensörleri ile sabit 50Hz hızında 3 eksenli doğrusal ivme ile 3 eksenli açısal hız kaydedilmiştir. 2.56 saniyelik sabit çerçeve genişliğinden elde edilen veriler örnekleme işleminden geçirilmiştir. Zaman (t) domaininde X,Y ve Z eksenlerinde sensörlerden yakalanan veriler Çizelge 3.2’de gösterilmiştir.

Çizelge 3.2. Sensörlerden elde edilen veriler

Sensör Verileri	
tBodyAcc-XYZ	tGravityAcc-XYZ
tBodyAccJerk-XYZ	tBodyGyro-XYZ
tBodyGyroJerk-XYZ	tBodyAccMag
tGravityAccMag	tBodyAccJerkMag
tBodyGyroMag	tBodyGyroJerkMag
fBodyAcc-XYZ	fBodyAccJerk-XYZ
fBodyGyro-XYZ	fBodyAccMag
fBodyAccJerkMag	fBodyGyroMag
fBodyGyroJerkMag	

Vücut hareketleri ve yer çekimi bileşenleri içeren ivmeölçer sinyali bir alçak geçirgen filtre kullanarak vücut ivmesi ve yer çekimi olmak üzere ayrılmıştır. Yer çekimi sinyalleri düşük frekans değerlerine sahip olduğu için 0,3Hz ‘lik filtre kullanılmıştır. Etiket verileri ise

manuel video kaydı ile elde edilmiştir. Her uygulama sırasında veri seti random olarak sırasıyla %90, %80, %70 eğitim ve %10, %20, %30 test olmak üzere ikiye bölünmüştür. Sensör verilerinden özellik çıkarmada kullanılan fonksiyonlar Çizelge 3.3'te gösterilmiştir.

Çizelge 3.3. Özellik çıkarmada kullanılan fonksiyonlar

Özellik Çıkarım Fonksiyonları	
mean()	Ortalama değer
std()	Standart sapma
mad()	Medyan mutlak sapma
max()	Dizideki en büyük değer
min()	Dizideki en küçük değer
sma()	Sinyal büyüklüğü alanı
energy()	Enerji ölçüsü. Karelerin toplamının değer sayısına bölümü.
iqr()	Çeyrekler arası aralık
entropy()	Sinyal entropisi
arCoeff()	Burg sırası 4'e eşit olan otoregresyon katsayıları
correlation()	İki sinyal arasındaki korelasyon katsayısı
maxInds()	En büyük öneme sahip frekans bileşeninin indeksi
meanFreq()	Ortalama bir frekans elde etmek için frekans bileşenlerinin ağırlıklı ortalaması
skewness()	Frekans alanı sinyalinin eğriliği
kurtosis()	Frekans alanı sinyalinin basıklığı
bandsEnergy()	Her pencerenin FFT değerinin bir frekans aralığının enerjisi
angle()	Vektörler arası açı

3.2 Veri Ön İşleme

Sensör verilerini tutarlı hale getirmek amacıyla gürültü filtreleri uygulanarak veri ön işleme adımları gerçekleştirilmiştir. Veri seti içerisinde zaman ve frekans domain değişkenleri ile birlikte Fast Fourier Transform (FFT) vb. yöntemler kullanılarak elde edilmiş 561 özellik vektörü bulunmaktadır. Özellik vektörleri her bir örnekleme çerçevesinden zaman ve frekans alanlarından hesaplanarak elde edilmiştir.

Özellik değerleri $[-1,1]$ arasına normalize edilmiştir. Normalizasyonun amacı veri seti içindeki değer aralıklarını bozmadan sayısal alanlardaki değerleri ortak bir ölçeğe uyarlamaktır. Her veri seti için normalizasyon gerekme de özellikler farklı aralıklara sahip olduğunda örneğin yaş ve gelir bilgisi vb. farklı aralıklardaki değerler için tercih edilmektedir. Gerektiğinde yapılmayan normalizasyon işlemi sonuçlarda olumsuz sapmaya neden olabilmektedir. Yapılan normalizasyon işleminin makine öğrenmesi modellerinin öğrenme hızı ve elde edilecek doğruluk oranlarına olumlu katkı yapacağı düşünülmektedir.

3.3 Jiroskop

Jiroskoplar içinde bulundukları çerçeveye döndüğünde açısal hızı algılayabilen sensörlerdir. Mekanik, optik, fiber optik, lazer gibi farklı sınıflarda birçok jiroskop çeşidi bulunmakla birlikte, navigasyon sistemleri, pusula, atalet ölçümü, mikro elektronik sistemler gibi karmaşık yapılarda kullanılabilirler. (Passaro, Cuccovillio, Vaiani, De Carlo, & Campanella, 2017)

Jiroskoplar hassas sensörlerdir ve dönme hareketini açısal momentum ilkelerine dayalı olarak ölçebilirler. Jiroskoptan elde edilen veriler, cihazın koordinat sistemlerinde gösterilen 3 farklı eksen etrafında ne kadar hızlı döndüğünü gösteren açısal hızdır. (Ming, 2013).

3.4 İvme Ölçer

Bir ivmeölçer sensörünün çalışma prensibi eylemsizlik kuvvetine dayanır. İvmeölçer sensörleri 3 ekseninde düşük gürültülü hızlanma ölçümüne olanak sağlamaktadırlar. Bu özellikleri sayesinde akıllı telefonlarda, tabletlerde, bilgisayarlarda, robotik sistemlerde ve sensör ağlarında eğim, hareket ve titreşimi algılayabilirler. (Hyunh & Yoo, 2016)

Temelde ivmeölçer sensörleri, kapasitif ve dirençli değişiklikler yoluyla ivmeyi elektrik sinyallerine dönüştürürler. Bu durumda bir ivmeölçer tasarımı yapabilmek için elektrik alanı alt yapısı gerekmektedir. Piezodirençli, kapasitif, optik olmak üzere farklı sınıflarda ivmeölçer sensörlerine rastlanmaktadır. (Kesilmiş, 2011)

(Kuzu, 2006)' ya göre ivme ölçerlerin ortak yönü uyarılan dahili ispat kümeleri kullanmalarıdır. Buna rağmen, dahili ispat kütle ivmelerini ölçmeyi sağlayan algılama mekanizmaları farklıdır. Kapasitif ivmeölçerler, denge kütlelerinin hareketini çıkarmak için kullanılan bir diferansiyel kapasitör kullanırken, piezodirençli ivmeölçerler genellikle ivmeölçerin yay kısmındaki gerilimi kullanırlar.

3.5 Analiz Yöntemleri

Yapılan uygulamalarda kullanılan veri seti, 4 farklı makine öğrenmesi algoritması ile bir derin öğrenme algoritması kullanan birbirinden farklı modellerle analiz edilmiş olup,

modellerin eğitiminde ve test işlemlerinde cross validation (çapraz doğrulama) yapılarak sonuçlarda tutarlılık kontrolü yapılmıştır.

Veriler modellere.xlsx formatında sunulmuş olup, eğitim ve test olmak üzere iki ayrı dosyadan oluşmaktadır.

Elde edilen sonuçların değerlendirilmesinde karmaşıklık matrisi, doğruluk metrikleri ve roc eğrilerinden yararlanılmıştır.

3.5.1 Makine öğrenmesi

Genel bir süreç olarak öğrenme, yeni bilgileri edinmek veya mevcut davranışları, değerleri, bilgileri, becerileri veya tercihleri değiştirmekle ilgilidir. Davranışçılık, Bilişselcilik, Yapılandırmacılık, Deneyimcilik ve Sosyal Öğrenme, kişisel öğrenme teorisini, yani insanların nasıl öğrendiğini tanımlar. Makineler ise insanlardaki doğal süreçlerin aksine verilere güvenir ve deneyimlerinden öğrenir. Temel olarak makine öğrenmesi bilgisayarların kendi başlarına düşünmelerini ve öğrenmelerini sağlayan bir yapay zeka kategorisidir. (Alzubi, Nayyar, & Kumar, 2018)

(Görmez, 2021)' e göre makine öğrenmesi, yazılım temelli öğrenmeyi sağlayan bilgisayar sistemleridir. Makine öğrenmesinde performans artışı öğrenme kriteri olarak kabul edilmeli ve bunun için veri setleri gerekmektedir. Öğrenme modelleri performansı artırmak amacıyla farklı parametreler kullanırken, veri setleri içerisinde seçilecek anlamlı bilgiler öğrenmeye katkı sağlayacaktır.

(Paksoy, 2021) yüksek lisans tezinde, literatür incelendiğinde makine öğrenmesi birçok yöntem içermekle birlikte üç ana başlık altında toplanabileceğini belirtmiştir. Bu yöntemler kullanılmadan önce test hatasından anlam çıkarabilmek amacıyla belirli oranlarda test ve eğitim verileri olmak üzere bölünür. Eğitim ve test verileri bir bağımsız değişken ve bir bağımlı değişkenden oluşabileceği gibi, birden fazla bağımsız değişken içeren bir matris de oluşturabilir. Makine öğrenmesi yöntemleri genel olarak denetimli öğrenme, denetimsiz öğrenme ve yarı denetimli öğrenme olarak sınıflandırılmaktadır.

Denetimli öğrenme bir öğretmen aracılığıyla öğrenmeye benzetilebilir. Denetimli öğrenmede model, girdileri ve çıktıları gözlemler ve aralarındaki ilişkiyi saptamaya çalışır. X_i giriş değerleri olarak adlandırılırsa, model bu girdilere karşılık $f(x_i)$ çıktıları üretecektir. Denetimli öğrenmede model, gerçek çıktılar ve üretilen çıktılar arasındaki

farklılıkları gözlemler. Bununla birlikte $y_i - f(x_i)$ sonucuna karşı girdi çıktı ilişkisini değiştirebilme özelliğine sahiptir.

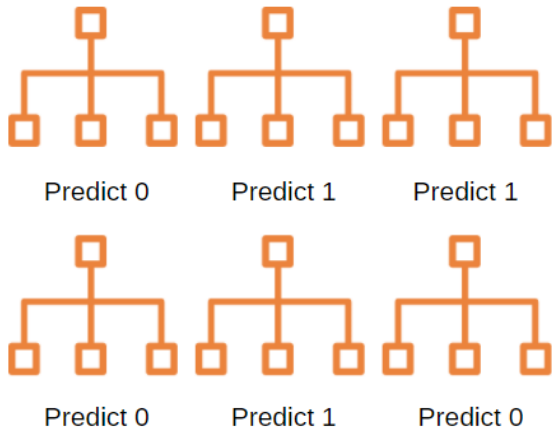
Denetimli öğrenme, girdileri ve çıktıları belirli örneklerden öğrenme olarak da adlandırılabilir. Süreç sonunda modelin, daha sonra karşılaşacağı girdiler için gerçeğe yakın sınıflandırma sonuçları üreteceği beklenmektedir. Aynı zamanda denetimli öğrenme çoğunlukla sınıflandırma problemlerinin çözümünde kullanılmaktadır. (Hastie, Tibshirani, & Friedman, 2008)

Denetimsiz öğrenmede amaç, her giriş değeri için doğru cevaplar veya hata oranları sunan bir örnek veya öğretmen olmaksızın olasılık yoğunluğunun özelliklerini doğrudan çıkarmaktır. Özellik çıkarımı bir sınıflandırma sonucu olarak değil benzer özelliklerin kümelenmesi sonucu olarak gözlemlenir. Denetimsiz öğrenmede seçilen model, giriş değişkenleri arasındaki ilişkiler ve bunların daha küçük bir gizli değişkenler kümesinin işlevleri olarak kabul edilip edilemeyeceği hakkında bilgi sağlar. (Hastie, Tibshirani, & Friedman, 2008)

(Aydemir, 2019)' a göre yarı denetimli öğrenme iki aşamadan oluşur. İlk aşamada etiketli veriler seçilir ve buna bağlı olarak ikinci aşamada etiketlenecek veriler hedeflenir. Yarı denetimli öğrenmede büyük veri kütlelerini etiketleme problemini çözmek amacıyla az miktardaki etiketli veriden yola çıkarak yüksek performans elde edebilmek amaçlanmıştır. Bu çalışmada denetimli öğrenme kullanan sınıflandırıcı algoritmalar kullanılmış ve aşağıda algoritmalar tanıtılmıştır.

3.5.2 Rastgele orman sınıflandırıcı (random forest)

Rastgele ormanlar, her ağacın bağımsız olarak örneklenen rastgele bir vektörün değerlerine bağlı olduğu ve ormandaki tüm ağaçlar için aynı dağılıma sahip olduğu, ağaç tahmin edicilerin bir kombinasyonudur. Genelleme hatası orman ne kadar geniş olursa olsun ağaç sayısı ile sınırlıdır. Bir orman sınıflandırıcısının genelleme hatası, ormandaki ağaçların gücüne ve aralarındaki korelasyona bağlıdır. Her bir ağaç bağımsız vektör girişleriyle en popüler sınıfa oy verir. (Breiman, 2001)



Şekil 3.1. Rastgele orman sınıflandırıcı

Rastgele orman adından da anlaşılacağı üzere birlikte çalışan çok sayıda bireysel karar ağacından oluşmaktadır. Şekil 3.1’de görüldüğü gibi, ormandaki her bir ağaç bir sınıf tahmini verir ve en çok oyu alan sınıf modelin tahmini olarak sunulur. (Yiu, 2019)

Karar ağaçlarında dallanmanın nasıl gerçekleşeceğini belirlemek amacıyla Gini veya Entropy kullanılabilir. (Schott, Random Forest Algorithm for Machine Learning, 2019)

$$\text{Gini} = 1 - \sum_{i=1}^c (p_i)^2 \quad (3.1)$$

Eş. 3.1’de p_i veri setinde gözlemlenen sınıfın göreceli frekansını c ise sınıf sayısını temsil etmektedir.

$$\text{Entropy} = \sum_{i=1}^c -p_i * \log_2(p_i) \quad (3.2)$$

Eş. 3.2’de gösterilen Entropy, düğümlerin nasıl dallanacağına karar verirken belirli bir sonucun olasılığını kullanmaktadır.

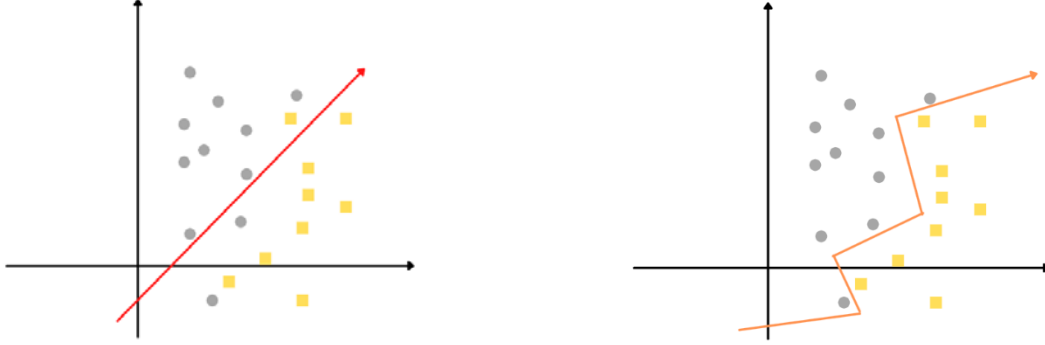
3.5.3 Destek vektör makinesi (support vector machine)

Destek vektör makineleri örüntü sınıflandırma problemleri için kullanılmakta ve başarılı sonuçlar elde edilmektedir. Destek vektörü yaklaşımını belirli bir probleme uygulamak, problem tanıma ve onunla ilgili tasarıma dayalı bir dizi sorunun çözülmesini içermektedir.

Uygulamada karşılaşılan zorluklardan biri uygun bir çekirdek seçmektir. Gauss ve polinom çekirdeği gibi seçenekler bulunmakla birlikte girdilerin ayrık yapılarda olması durumunda daha ayrıntılı çekirdeklere ihtiyaç duyulacaktır. (Jakkula, 2011)

SVM, ayırıcı bir hiper düzlem ile tanımlanan ayırt edici bir sınıflandırıcıdır. Yani, etiketli eğitim verileri verildiğinde, algoritma yeni örnekleri kategorize eden optimal bir hiper düzlem üretir.

(Patel, 2017)' e göre iki boyutlu uzayda bu düzlem, bir düzlemi her bir sınıfın her iki tarafta da bulunduğu iki parçaya bölen bir çizgidir.



Şekil 3.2. SVM parametreye bağlı düzlem oluşumu

Problemin doğrusal cebir kullanılarak dönüştürülmesiyle, doğrusal SVM’de hiper düzlemin öğrenilmesi sağlanır. Öğrenmeye bağlı olarak düzlemdeki değişim Şekil 3.2’de gösterilmiştir. Bu noktada ‘kernel’ parametresi seçimi yapılmaktadır. Doğrusal ‘kernel’ için girdi(x) ile her destek vektörü(x_i) arasındaki iç çarpımı kullanan yeni bir girdi için tahmin denklemi Eş. 3.3’te gösterilmiştir.

$$f(x) = B(0) + \text{sum}(a_i * (x, x_i)) \quad (3.3)$$

Eş. 3.3’ te verilen $B(0)$ ve a_i katsayıları öğrenme algoritması tarafından eğitim verilerinden tahmin edilmelidir.

$$K(x, x_i) = 1 + \text{sum}(x * x_i)^d \quad (3.4)$$

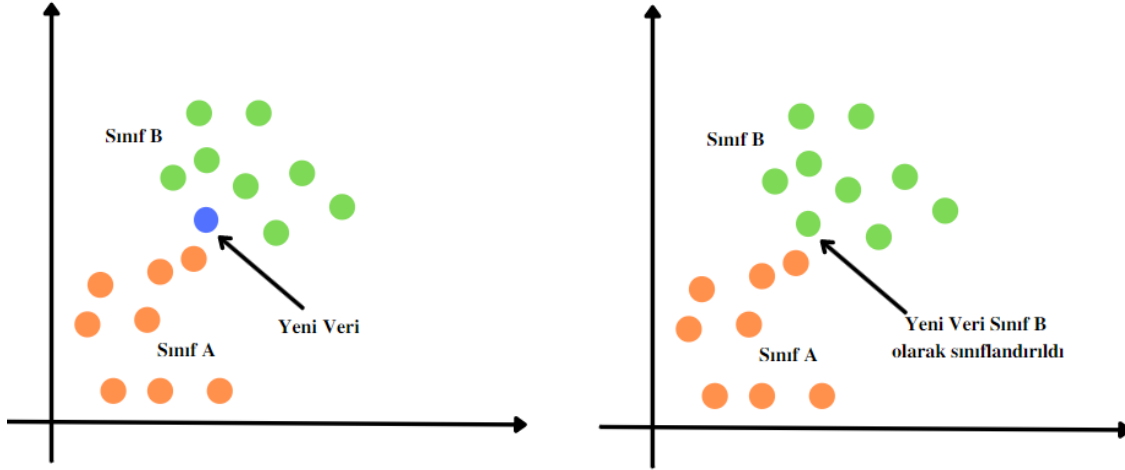
$$K(x, x_i) = \exp(-\text{gamma} * \text{sum}(x - x_i)^2) \quad (3.5)$$

Polynomial kernel için Eş. 3.4, exponential kernel için Eş. 3.5 kullanılmaktadır.

3.5.4 En yakın komşuluk (k-nearest neighbour)

K-en yakın komşuluk algoritması, hem sınıflandırma hem de regresyon problemlerinin çözümü için kullanılabilen bir denetimli makine öğrenmesi algoritması türüdür. Daha çok endüstride sınıflandırma problemlerinin çözümü için kullanılmaktadır.

KNN algoritması, yeni durum veya veriler için mevcut durumlar arasındaki benzerliği varsayar ve yeni durumu mevcut kategorilere en çok benzeyen kategoriye koyar. KNN tembel öğrenen algoritma olarak da adlandırılmaktadır. Çünkü eğitim kümesinden hemen öğrenmez, bunun yerine veri kümesini depolar ve sınıflandırma anında veri kümesinde işlem gerçekleştirir. (Datasciencelovers, 2020)



Şekil 3.3. K-en yakın komşuluk sınıflandırması

En yakın komşuluk sınıflandırması görsel olarak Şekil 3.3'te gösterilmiştir. KNN' de sınıflandırma yapılırken veri noktaları arasındaki uzaklık Eş. 3.6'da gösterilen Euclidean yöntemi ile hesaplanmaktadır. (Schott, K-Nearest Neighbors (KNN) Algorithm for Machine Learning, 2019)

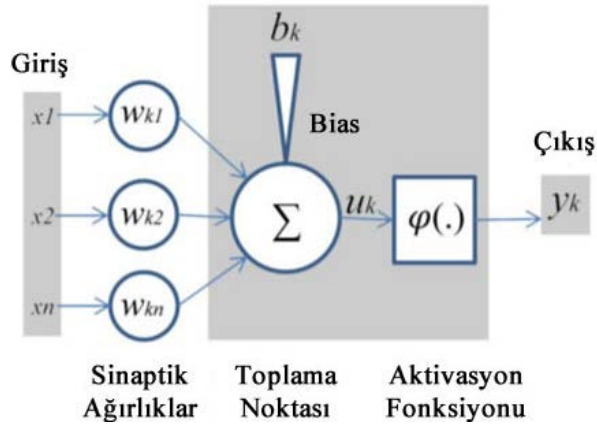
$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (3.6)$$

3.5.4 Çok katmanlı algılayıcı (multi layer perceptron)

Yapay Sinir Ağı (YSA) algoritmaları, anlama, öğrenme, problem çözme ve karar alma gibi insan beynine benzer işlevleri yerine getiren bir metodoloji kullanarak ilgi alanlarını sınıflandırmada kullanılmaktadır. Genel olarak sinir ağları, beynin belli bir görevi veya ilgilenilen işlevi yerine getirme şeklini modellemek için tasarlanmış bir makinedir. Şekil 3.4 'te görüldüğü üzere bir yapay sinir ağı mimarisi üç birimden oluşmaktadır (Taravat, Proud, Peronaci, Frate, & Oppelt, 2015).

İlk katman girdi katmanıdır ve düğümlerin sayısı girdi parametreleri tarafından belirlenir. Son katman çıktı katmanıdır ve düğüm sayısı beklenen çıktı tarafından belirlenir. Giriş ve çıkış katmanları arasında kalan katmanlar gizli katmanlar olarak adlandırılır. Gizli katmanlar veri işlemlerde çoğunlukla doğrusal olmayan aktivasyon fonksiyonları kullanır.



Şekil 3.4. Doğrusal olmayan bir yapay sinir ağı modeli

Şekil 3.4'te görülen b_k bias değeri aktivasyon fonksiyonunun etkisini artırma veya azaltma etkisine sahiptir. Bir k nöronu Eş. 3.7 ve Eş. 3.8'de gösterildiği gibi tanımlanabilir.

$$u_k = \sum_{i=1} w_{ki} x_i \quad (3.7)$$

$$y_k = \varphi(u_k + b_k) \quad (3.8)$$

$x_1, \dots, x_n$ giriş parametreleri, $w_{k1}, \dots, w_{kn}$ k nöronuna ait ağırlıkları, u_k giriş parametrelerinden kaynaklanan doğrusal toplamı, b_k bias değerini, $\varphi(.)$ aktivasyon fonksiyonunu ve y_k nöronun çıkış sinyalini temsil etmektedir. MLP modeli ileri beslemeli bir yapay sinir ağı sınıflandırıcısıdır.

MLP algoritmalarının uygulanmasında, çoğunlukla parametrelerin seçilmesi ve ağ konfigürasyonu ile ilgili pratik problemlerle karşılaşmaktadır. Örneğin algoritmanın öğrenme hızı düşük seçildiğinde sonuca yakınsama daha yavaş olurken, çok yüksek hız başarısız sonuçlara neden olabilir. Karşılaşılabilecek başka bir sorun ise hatalı özellik seçiminden kaynaklanan aşırı öğrenme (*overfitting*) durumudur. (Marius, Mastorakis, & Balas, 2009)

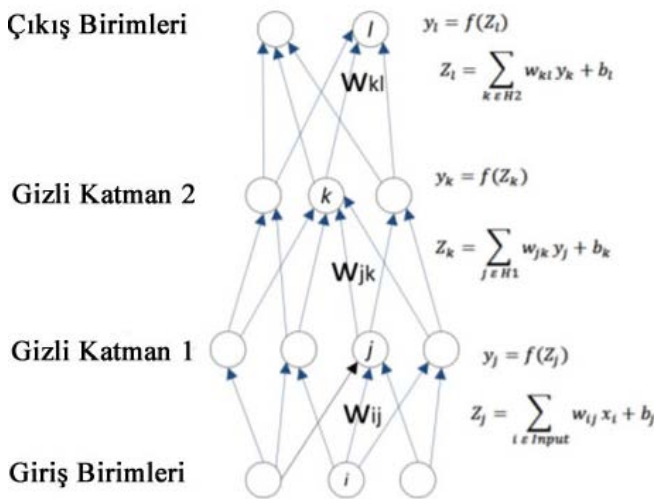
3.5.5 Derin öğrenme (deep learning)

Derin öğrenme, veriler üzerinde üst düzey soyutlama gerçekleştirebilen, gizli katmanlarla birlikte karmaşık ve doğrusal olmayan yapılardan oluşan bir öğrenme algoritması türüdür. Derin öğrenme mimarisi sığ yapay sinir ağları gibi ara katmanlara sahip geri beslemeli ağlardır. Bununla birlikte, derin öğrenme modellerinde daha fazla katmanla birlikte daha yüksek düzeyde soyutlama gerçekleştirilirken modelin yeteneği de gelişmektedir. (Mohammed, 2021)

Derin öğrenme mimarileri sınıflandırma problemlerinin çözümü, görüntü sınıflandırma, zaman serilerinin tahmini, doğal dil işleme gibi işlevleri yerine getirir. Derin öğrenme modelleri ile denetimli, denetimsiz veya yarı denetimli öğrenme gerçekleştirilebilir. Şekil 3.5’ te görüldüğü üzere model giriş katmanı, gizli katmanlar ve çıkış katmanından oluşmaktadır. (Shrestha & Mahmood, 2019)

Her bir nöronun çıkışa etki eden bir ağırlık w ve bias b değeri bulunmaktadır. Geri beslemeli ağlarda tahmin edilen sonuç ile gerçek sonuç arasındaki hata hesaplanır ve hata geriye yayılarak nöronların ağırlıkları güncellenir, optimum değer elde edilene kadar işlem tekrar eder.

Ağırlık parametreleri güncellenirken, $w = w - \alpha dW$ bias parametreleri güncellenirken $b = b - \alpha db$ formülleri kullanılır. dW ve db değerleri zincir kuralı ile hesaplanırken kayıp fonksiyonunun W ve b ’ ye göre kısmi türevlerini ifade eder.



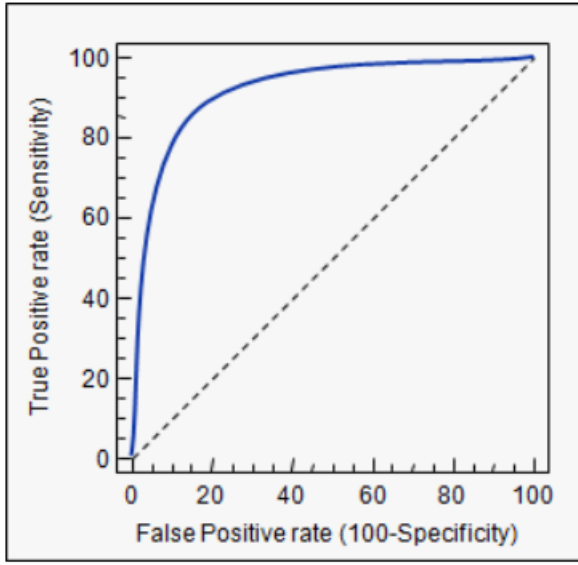
Şekil 3.5. Derin öğrenme model mimarisi

Şekil 3.5’te örnek bir derin öğrenme mimarisi gösterilmiştir. Derin öğrenmenin geleneksel makine öğrenmesi yöntemlerine göre ana avantajı, daha büyük veri kümelerinde daha iyi öğrenme performansı göstermesidir. Bu durum kullanılan verinin karakteristiği ile yakından ilişkilidir. Derin öğrenme ağlarına CNN (Evrişimli Sinir Ağları), RNN (Tekrarlayan Sinir Ağları), LSTM (Uzun-Kısa Dönem Hafıza) örnek gösterilebilir. CNN ağları daha çok görüntü işlemede kullanılırken, RNN ve LSTM ağları zaman serilerinin tahmini için tercih edilmektedir. (Sarker, 2021)

3.5.6 ROC eğrileri ve değerlendirme metrikleri

Bir Alıcı Çalışma Karakteristiği (ROC) eğrisi, çeşitli eşik değerlerde gerçek pozitif oran (TPR) ile yanlış pozitif oran (FPR) arasında çizilerek oluşturulur.

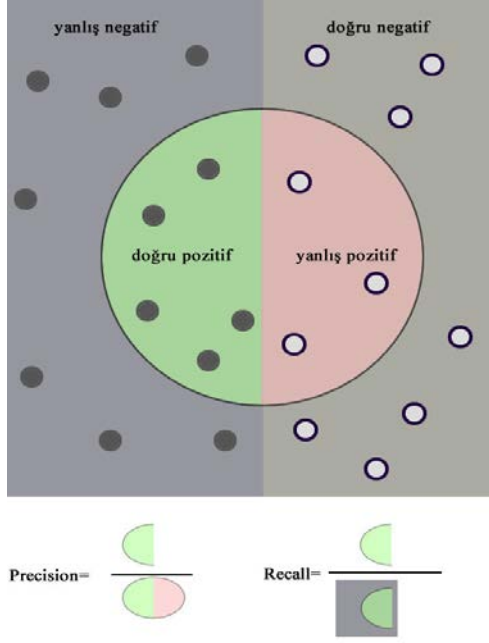
Makul doğrulukta bir test verisinin grafikteki sol üst üçgende, referans çizgisinin üzerinde bir ROC eğrisine sahip olması gerekmektedir. ROC eğrilerinde altta kalan alan Area Under The ROC Curve (AUC) test verisi hakkında fikir veren genel bir ölçüt olarak kabul edilmektedir. (Hoo, Candlish, & Teare, 2017)



Şekil 3.6. Roc eğrisi

ROC eğrileri ikili sınıflandırmada kullanıldığı gibi çoklu sınıflandırma problemleri için de çizilebilirler. Bir sınıf için çizilmiş ROC eğrisi Şekil 3.6'da gösterilmiştir (veribilimcisi, 2021).

Aktivite sınıflandırma sonuçları değerlendirilirken Precision, Recall, F1-Score ve Accuracy metriklerinden de yararlanılmıştır.



Şekil 3.7. Precision ve recall metrikleri örnek gösterim

Precision ve Recall metriklerinin formüle edilmesi Şekil 3.7’de gösterilmiştir. Örneğin kanser hastalığına sahip kişiler için tahmin yapılırken gerçekten kanser hastası olan ve hasta olarak tahmin edilenlerin sayısı doğru pozitif (true positive- TP), kanser hastası olmadığı halde hasta olarak tahmin edilenlerin sayısı yanlış pozitif (false positive-FP) olarak adlandırılır. Kanser hastası olduğu halde hasta olarak tahmin edilmeyenlerin sayısı yanlış negatif (false negative-FN), kanser hastası olmadığı halde hasta olmadığı tahmin edilenlerin sayısı doğru negatif (true negative-TN) olarak adlandırılır.

Precision metriği, Eş. 3.9’da gösterilen doğru pozitif tahminlerin, doğru pozitif ve yanlış pozitif tahminlerin toplamına oranı olarak adlandırılır.

$$P = \frac{TP}{TP + FP} \quad (3.9)$$

Recall metriği, Eş. 3.10’da gösterilen test verisinde tahmin edilen doğru pozitif verilerin, test verisinde tahmin edilen doğru pozitif ve yanlış negatif verilerin toplamına oranıdır.

$$R = \frac{TP}{TP + FN} \quad (3.10)$$

F1-Score metriği, Eş. 3.11’de gösterilmiştir ve Precision ve Recall metriklerinin harmonik ortalamasıdır.

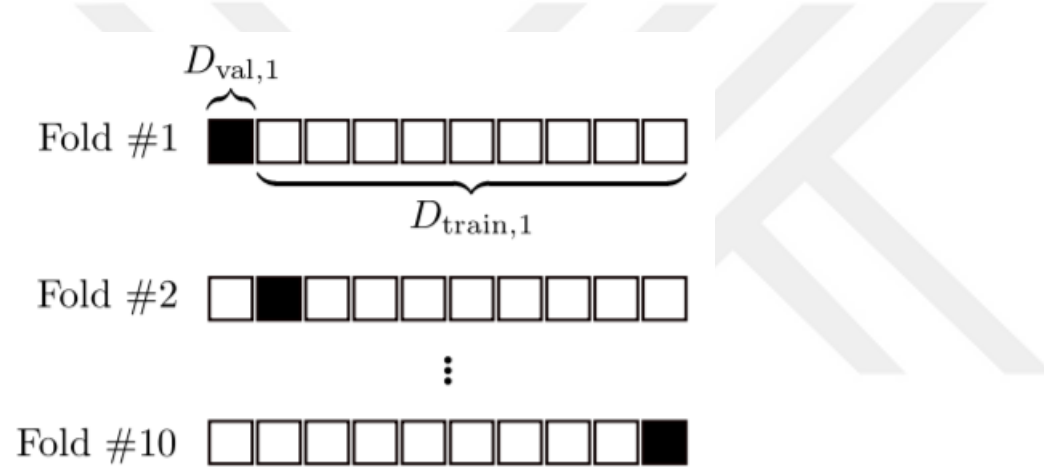
$$F = \frac{2PR}{P + R} \quad (3.11)$$

Accuracy metriği test verisinde tahmin edilen doğru pozitif ve doğru negatif tahminlerin toplam tahmin sayısına oranıdır. Eş. 3.12’de gösterilen Accuray değeri 1’e yaklaştıkça modelin doğruluk başarı oranı artarken, 0’a yaklaştıkça azalmaktadır.

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.12)$$

3.5.7 Çapraz doğrulama (cross validation)

Çapraz doğrulama, modellerin gerçek tahmin hatasını tahmin etmek ve model parametrelerini ayarlamak için en yaygın olarak veri yeniden örnekleme yöntemlerinden biridir. Tahmine dayalı modellerin genelleme yeteneğini değerlendirmek için tercih edilir. Çapraz doğrulama ile tekrarlanan alt örneklemeyle dayalı model değerlendirmesi, bir öğrenme fonksiyonunun birkaç veri alt kümesine uygulanarak yapılabilmektedir. (Berrar, 2019)



Şekil 3.8. Eğitim verileri üzerinde çapraz doğrulama

Şekil 3.8’de gösterildiği üzere eğitim verileri 10 alt parçaya bölünerek eğitim ve doğrulamada rastgele veri miktarları ile model eğitilebilmektedir (Berrar, 2019). Elde edilen sonuçların tutarlılığının izlenmesi açısından etkin olduğu söylenebilir.

3.5.8 Hata matrisi (confusion matrix)

Hata matrisi makine öğrenmesi algoritmalarından elde edilen sınıflandırma problemlerine ait sonuçları değerlendirirken yararlanılan görsel bir matristir. Sınıflandırma sonucu mutlaka dört farklı değerlendirme içerir. Bunlar,

- Gerçekte doğru olanı doğru olarak tahmin etmek (True Positive-TP)
- Gerçekte yanlış olanı yanlış olarak tahmin etmek (True Negative-TN)
- Gerçekte doğru olanı yanlış olarak tahmin etmek (False Positive-FP)
- Gerçekte yanlış olanı doğru olarak tahmin etmek (False Negative-FN)

şeklindedir.

		TAHMİN		
		YOK	VAR	TOPLAM
GERÇEK	YOK	T N 100	F P 20	120
	VAR	F N 10	T P 200	210
	TOPLAM	110	220	

Şekil 3.9. Gerçek ve tahminleri içeren bir hata matrisi

Şekil 3.9’da örnek bir hata matrisi gösterilmiştir (Şirin, 2017). Kullanılan algoritmanın güncel gerçek değerleri doğru tahmin edip etmediğini görmemiz açısından tercih edilmektedir.

4. BULGULAR VE TARTIŞMA

Bu bölümde Bölüm 3'te yer verilen makine öğrenmesi algoritmaları kullanılarak elde edilen sonuçlar metriklerle birlikte tablolar halinde paylaşılmıştır. ROC eğrilerinin grafiksel gösterimi yapılarak veri seti içindeki sınıfların analizi yapılmıştır. Kullanılan makine öğrenmesi yöntemleri ile elde edilen sonuçlara ait karmaşıklık matrislerine de yer verilmiştir.

Tüm modellerde veriler eğitim amaçlı kullanılmadan önce sklearn kütüphanesinde yer alan ExtraTreesClassifier yöntemi ile özellik seçimi (feature selection) yapılmıştır. Seçilecek özellikler için eşik değeri parametresi (threshold) 0.001 olarak alınmış, sonuca etki eden ve eşik değerini aşan özellikler seçilmiştir. Veri seti ile daha önce yapılan çalışmalardan elde edilen sonuçlar ve benzer çalışmalardan elde edilen sonuçlarla karşılaştırma yapılmış ve tartışılmıştır.

Modellerin eğitimi sırasında tüm modellerde sonuçların tutarlılığının izlenmesi amacıyla çapraz doğrulama (cross validation) uygulanmıştır. Çapraz doğrulama ile eğitim verileri 10 alt parçaya bölünerek her bir parça ile eğitim ve doğrulama işlemi gerçekleştirilmiş ve elde edilen tahmin sonuçlarının birbirine yakın ve tutarlı olduğu gözlemlenmiştir.

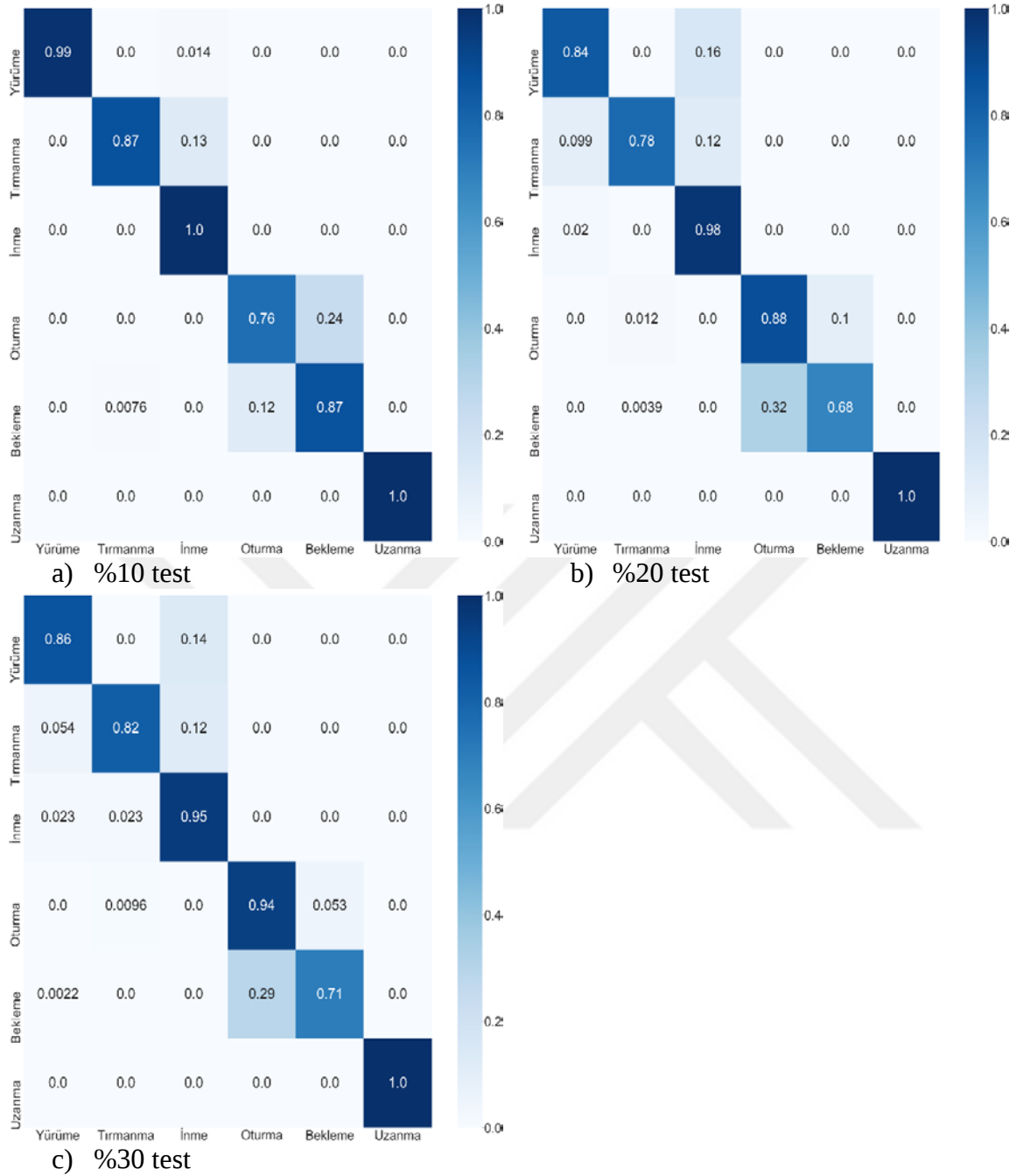
Veriler sırasıyla %90, %80, %70 eğitim ve %10, %20, %30 test olmak üzere bölünmüş ve elde edilen sonuçlar karşılaştırılmıştır.

4.1. KNN Sınıflandırıcısı ile Elde Edilen Sonuçlar

K-en yakın komşuluk algoritmasıyla yapılan uygulamada farklı komşuluk değerleri ile denemeler yapılmış ve en iyi sonuç veren değer seçilmiştir. Hata matrisleri ile yüzdelik gösterim yapılmış ve modelde elde edilen metriklerle roc eğrileri gösterilmiştir.

4.1.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi

KNN algoritması ile yapılan sınıflandırmada, komşuluk değeri 4 olarak alınmıştır. Elde edilen hata matrislerinde yüzdelik değerler Şekil 4.1'de gösterilmiştir.



Şekil 4.1. KNN algoritması hata matrisleri a) %10 test b) %20 test c) %30 test

Şekil 4.1'e göre, veriler %90 eğitim %10 test olarak bölündüğünde, inme ve uzanma aktivitelerinde sınıflandırma %100 başarılıdır. Yürüme aktivitesinde başarı oranı %99 iken aktivitenin %0,14' ü inme olarak sınıflandırılmıştır. Oturma aktivitesinin %24' ü bekleme olarak tahmin edilirken başarı oranı %76' dır. Tırmanma ve Bekleme aktivitelerinde başarı %87 iken tırmanma aktivitesinin %13' ü inme, bekleme aktivitesinin %12'si oturma ve %0,076' sı tırmanma olarak sınıflandırılmıştır.

Veri seti %80 eğitim %20 test olmak üzere bölündüğünde, uzmanma aktivitesinde sınıflandırma %100 başarılıdır. İnme aktivitesinde sınıflandırma başarısı %98 iken aktivitenin %2' si yürüme olarak tahmin edilmiştir. Oturma aktivitesinin %10' u bekleme, %1,2' si tırmanma aktivitesi olarak sınıflandırılırken başarı oranı %88' dir. Tırmanma aktivitesinin %0,99' u yürüme ve %12' si inme, yürüme aktivitesinin %16' sı tırmanma olarak belirlenmiştir. Tırmanma aktivitesinde sınıflandırma başarısı %78 iken yürüme aktivitesinde %84' dür. Bekleme aktivitesinin %32' si oturma %0,39' u tırmanma aktivitesi olarak tanımlanmıştır. Bekleme aktivitesinde başarı %68' dir.

Veri seti %70 eğitim ve %30 test olarak bölündüğünde uzmanma aktivitesinde sınıflandırma %100 başarılıdır. Yürüme aktivitesinin %14' ü inme aktivitesi olarak tahmin edilmiştir ve başarı oranı %86' dır. Tırman aktivitesinde başarı oranı %82 ve inme aktivitesinde %95 olarak görülmektedir. Tırmanma aktivitesinin %12' si inme, %5,4' ü ise yürüme aktivitesi olarak belirlenmiştir. İnme aktivitesinin %2,3' ü yürüme, %2,3' ü ise tırman aktivitesi olarak tahmin edilmiştir. Oturma aktivitesinde Başarı oranı %94' tür. Aktivitenin %5,3' ü bekleme, %0,96' sı tırmanma olarak tahmin edilmiştir.

Çizelge 4.1. KNN ile elde edilen hata matrisleri örneklem gösterimi

	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma
	%90 Test %10 Eğitim					
Yürüme	143	0	2	0	0	0
Tırmanma	0	96	14	0	0	0
İnme	0	0	102	0	0	0
Oturma	0	0	0	92	29	0
Bekleme	0	1	0	16	114	0
Uzanma	0	0	0	0	0	126
	%80 Eğitim %20 Test					
Yürüme	227	0	43	0	0	0
Tırmanma	22	174	27	0	0	0
İnme	4	0	193	0	0	0
Oturma	0	3	0	222	26	0
Bekleme	0	1	0	82	176	0
Uzanma	0	0	0	0	0	270
	%70 Eğitim %30 Test					
Yürüme	358	0	56	0	0	0
Tırmanma	21	318	48	0	0	0
İnme	8	8	326	0	0	0
Oturma	0	4	0	391	22	0
Bekleme	1	0	0	135	326	0
Uzanma	0	0	0	0	0	428

Çizelge 4.1’de gösterilen hata matrislerinde veriler %90 eğitim %10 test olarak bölündüğünde KNN modeli 102 inme aktivitesinin ve 126 uzanma aktivitesinin tamamını doğru tahmin etmiştir. 145 yürüme aktivitesinin 2’ si inme, 110 tırmanma aktivitesinin 14’ ü inme, 121 oturma aktivitesinin 29’ u bekleme, 131 bekleme aktivitesinin 16’ sı oturma ve 1’ i tırmanma olarak tahmin edilmiştir.

Veriler %80 eğitim %20 test olarak bölündüğünde, 270 uzanma aktivitesinin tamamı doğru tahmin edilmiştir. 270 yürüme aktivitesinin 43’ ü inme, 223 tırmanma aktivitesinin 22’ si yürüme ve 27’ si inme, 197 inme aktivitesinin 4’ ü yürüme, 251 oturma aktivitesinin 26’ sı bekleme, 3’ ü tırmanma olarak belirlenmiştir. 259 bekleme aktivitesinin 82’ si oturma ve 1’ i tırmanma aktivitesi olarak sınıflandırılmıştır.

Model %30 test verisiyle, 428 uzanma verilerinin tamamını doğru tahmin etmiştir. 414 yürüme aktivitesinin 56’ sı inme, 387 tırmanma aktivitesinin 21’ i yürüme ve 48’ i inme, 342 inme aktivitesinin, 8’ i yürüme, 8’ i tırmanma olarak tahmin edilmiştir. 417 oturma

aktivitesinin 22' si bekleme, 4' ü tırmanma olarak sınıflandırılmıştır. 462 bekleme aktivitesinin 135' i oturma ve 1' i yürüme aktivitesi olarak tahmin edilmiştir.

Üç farklı uygulamada en büyük sapmanın oturma ve bekleme aktivitelerinde olduğu görülmektedir. Oturma aktivitesinin bekleme aktivitesi olarak veya bekleme aktivitesinin oturma aktivitesi olarak sınıflandırılması, iki aktivitenin benzer şekilde hareketsizlik içermesinden kaynaklanıyor olabilir. Veri setleri içindeki oturma ve bekleme aktivitelerine ait sensör verilerinin değerlerinin birbirine yakın olduğu düşünülmektedir.

4.1.2 Test verisi ile metrikler

KNN modelinde elde edilen precision, recall, f1-score ve accuracy metrik değerleri ile test verilerine ait örneklem sayıları Çizelge 4.2'de gösterilmiştir.

Çizelge 4.2 KNN modeli ile elde edilen metrikler

	Precision %	Recall %	F1-Score %	Acc %	Örneklem Sayısı
	%90 Eğitim %10 Test				
Yürüme	1,00	0,99	0,99	0,92	145
Tırmanma	0,99	0,87	0,93		110
İnme	0,86	1,00	0,93		102
Oturma	0,85	0,76	0,80		121
Bekleme	0,80	0,87	0,83		131
Uzanma	1,00	1,00	1,00		126
	%80 Eğitim %20 Test				
Yürüme	0,90	0,84	0,87	0,86	270
Tırmanma	0,98	0,78	0,87		223
İnme	0,73	0,98	0,84		197
Oturma	0,73	0,88	0,80		251
Bekleme	0,87	0,68	0,76		259
Uzanma	1,00	1,00	1,00		270
	%70 Eğitim %30 Test				
Yürüme	0,92	0,86	0,89	0,88	414
Tırmanma	0,96	0,82	0,89		387
İnme	0,76	0,95	0,84		342
Oturma	0,74	0,94	0,83		417
Bekleme	0,94	0,71	0,80		462
Uzanma	1,00	1,00	1,00		428

%90 eğitim %10 test verisi ile Precision (hassaslık) metriğinde en yüksek hassaslık değerleri yürüme ve uzanma aktivitelerinde %100 olarak gerçekleşmiştir. En düşük

hassaslık değeri %80 ile bekleme aktivitesinde gerçekleşmiştir. Tırmanma aktivitesinde %99, inme aktivitesinde %86, oturma aktivitesinde %85 hassaslık değerleri elde edilmiştir. Hassaslık metriğinde, veri seti %80 eğitim %20 test olarak bölündüğünde uzanma aktivitesinde %100 değeri elde edilmiştir. Tırmanma aktivitesinde %98, yürüme aktivitesinde %90, bekleme aktivitesinde %87 değerleri elde edilirken en düşük değerler %73 ile inme ve oturma aktivitelerinde gerçekleşmiştir.

Veri seti %70 eğitim %30 test olarak bölündüğünde ise uzanma aktivitesinde %100 hassaslık elde edilmiştir. Tırmanma aktivitesinde %96, bekleme aktivitesinde %94, yürüme aktivitesinde %92, inme aktivitesinde %76 değerleri elde edilirken en düşük hassaslık değeri %74 ile oturma aktivitesinde görülmektedir.

Recall metriği her bir sınıf için yapılan tahminlerde doğruluk oranını göstermektedir. %90 eğitim %10 test verisi ile en yüksek doğruluk oranları %100 ile inme ve uzanma aktivitelerinde gerçekleşmiştir. En düşük doğruluk oranı %76 ile oturma aktivitesinde gerçekleşmiştir. Yürüme aktivitesinde %99, tırmanma ve bekleme aktivitelerinde %87 doğruluk oranları elde edilmiştir.

Veri seti %80 eğitim %20 test olarak bölündüğünde recall metriğinde en yüksek değerler %100 ile uzanma aktivitesinde görülmektedir. İnme aktivitesinde %98, yürüme aktivitesinde %84, oturma aktivitesinde %88, tırmanma aktivitesinde %78 değerleri elde edilmiştir. En düşük değer %68 ile bekleme aktivitesinde gerçekleşmiştir.

%70 eğitim %30 test verileri ile, uzanma aktivitesinde %100 değeri elde edilmiştir. İnme aktivitesinde %95, oturma aktivitesinde % 94, yürüme aktivitesinde %86, tırmanma aktivitesinde %82 değerleri elde edilmiştir. En düşük değer %71 ile bekleme aktivitesinde gerçekleşmiştir.

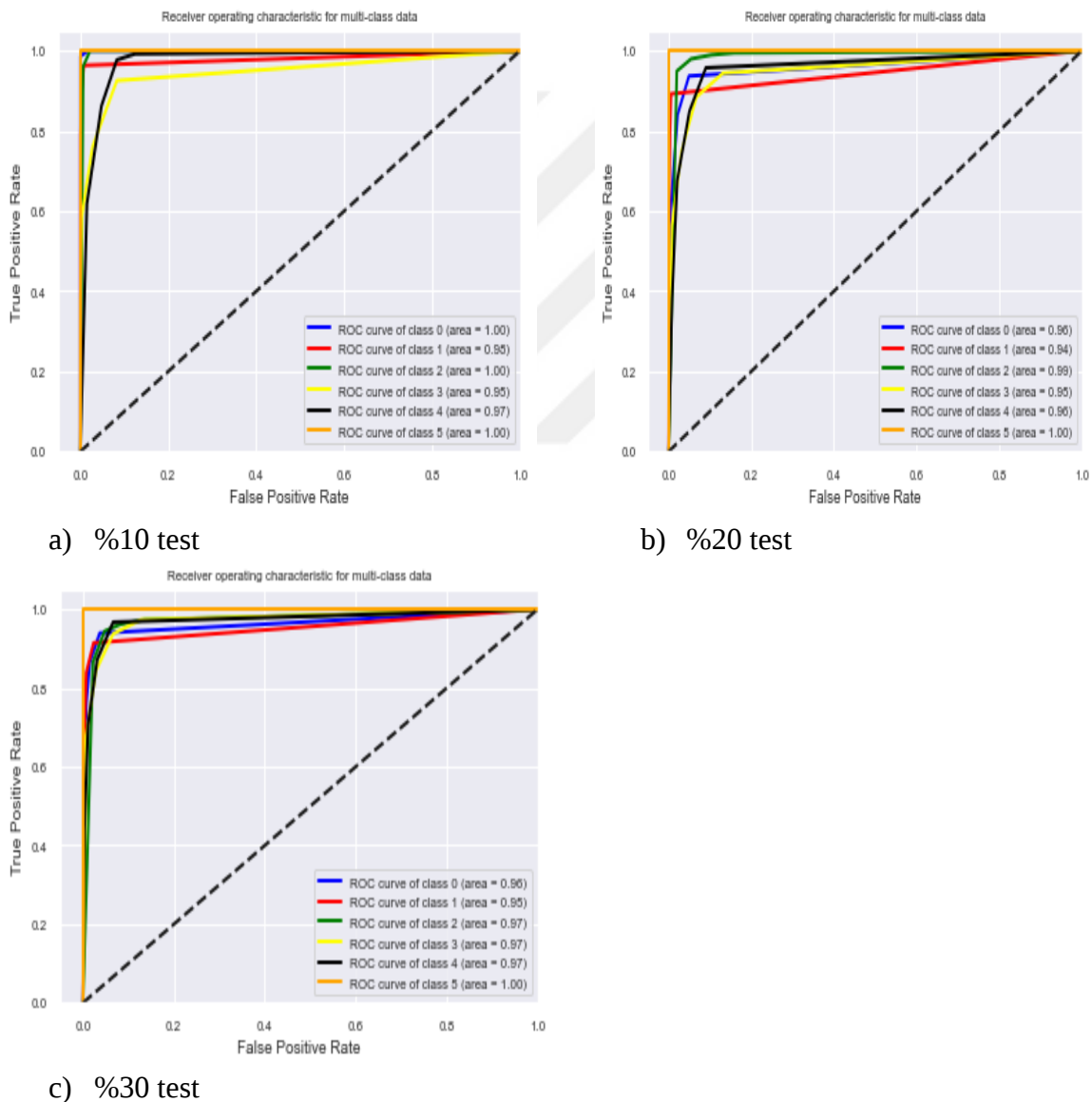
F1-score metriği precision ve recall metriklerinin bir ortalamasıdır. %10 test verisi ile en yüksek değer %100 ile uzanma aktivitesinde görülmektedir. Yürüme aktivitesinde %99, inme aktivitesinde %93 ve oturma aktivitesinde %80 değerleri elde edilmiştir. İkinci uygulamada %20 test verisi kullanarak elde edilen en yüksek değer %100 ile uzanma aktivitesinde görülmektedir. Yürüme ve tırmanma aktivitelerinde %87 iken en düşük değer, bekleme aktivitesinde %76' dır. %30 test verisinde uzanma aktivitesinde %100, yürüme ve tırmanma aktivitelerinde %89, inme aktivitesinde %84 ve en düşük değer %80 ile bekleme aktivesinde elde edilmiştir.

KNN algoritması ile oluşturulan modelin üç farklı uygulamada %10 test verisi ile %92, %20 test verisi ile %86, %30 test verisi ile %88 doğruluk oranları (accuracy) elde edilmiştir. Uygulamalarda sırasıyla tüm aktivitelerden toplam 735, 1470 ve 2450 test

örnekleme alınmış ve oluşan test verisinde sınıfların dağılımı örneklem sayısında gösterilmiştir. Sınıflara ait farklı test verileriyle elde edilen metriklerde farklılıklar gözlenmiştir. Test verisi arttıkça ortalama doğruluk oranı azalmıştır.

4.1.3 Test verisi ile roc eğrisi

KNN modelinde %10, %20 ve %30 oranlarında test verisiyle elde edilen ve test doğruluk oranlarını gösteren ROC eğrileri Şekil 4.2’de gösterilmiştir.



Şekil 4.2. KNN modeli ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test

Şekil 4.2’ye göre testlerin doğruluk oranı üç ayrı uygulamada, uzanma aktivitesinde %100 çıkarken, diğer aktivitelere ait verilerde farklı test verisi oranlarıyla % 94 ile %97 arasında

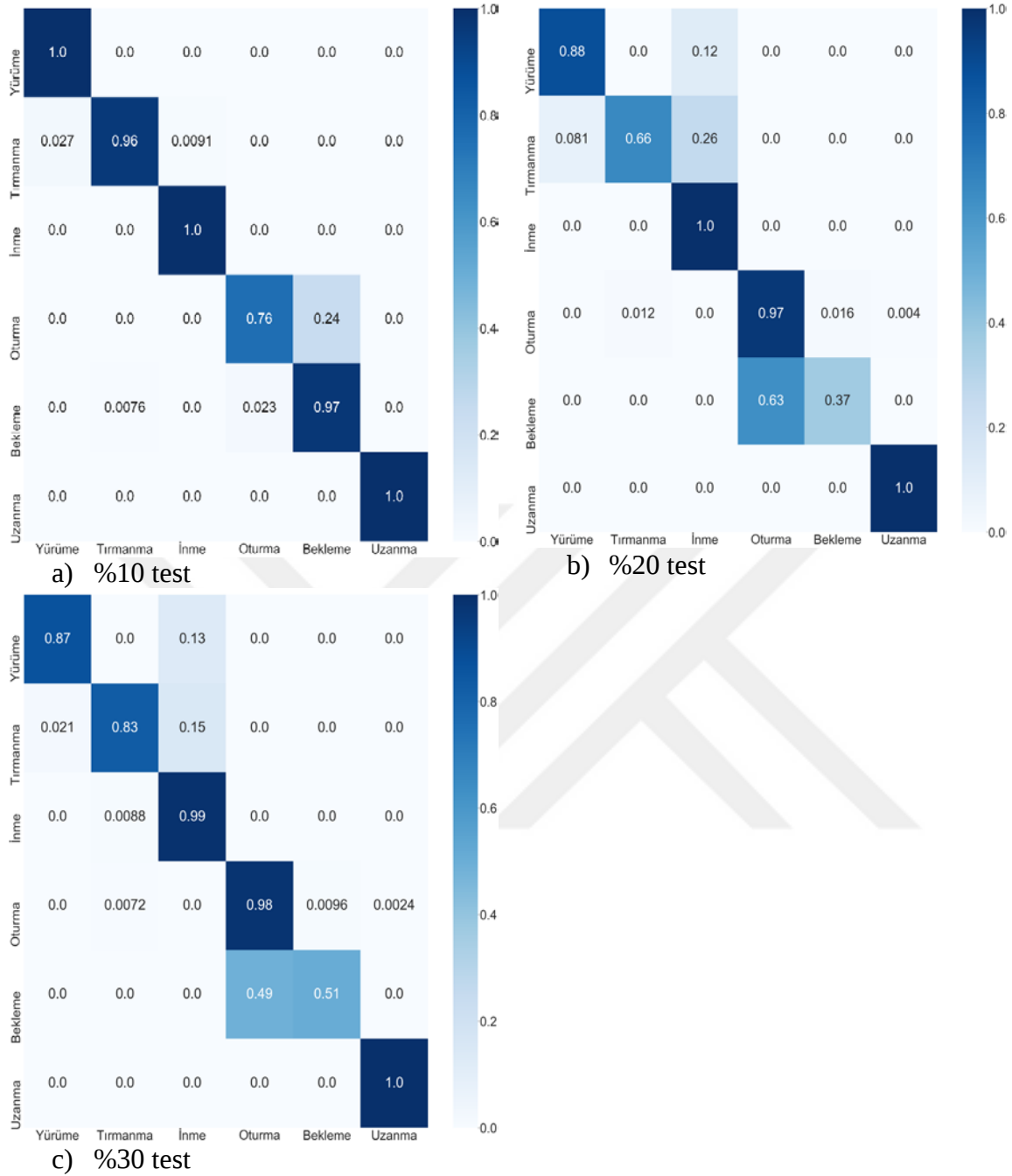
değişen değerler elde edilmiştir. Grafikte köşegen çizgisi üzerinde yer alan sınıflarda test verileri için değerlendirme yapılabilir. Sonuçlara göre test verilerinin makul olduğu söylenebilir.

4.2 SVM Sınıflandırıcısı ile Elde Edilen Sonuçlar

Destek vektör makinesi algoritması ile yapılan uygulamalarda farklı kernel parametreleri ile denemeler yapılmış ve en iyi sonuçları veren değerler seçilmiştir. Hata matrisleri ile yüzdelik gösterim yapılmış ve modelde elde edilen metriklerle roc eğrileri gösterilmiştir.

4.2.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi

Destek Vektör Makinesi (SVM) algoritması kullanılarak oluşturulan modelde, kernel parametresi 'rbf', gamma parametresi 'scale', ve degree parametresi 3 olarak seçilmiş ve eğitim yapılmıştır. Elde edilen hata matrislerinde yüzdelik değerler Şekil 4.3'te gösterilmiştir.



Şekil 4.3. SVM algoritması hata matrisleri a)%10 test b)%20 test c)%30 test

Matrislere göre, %10 test verisinde yürüme, inme ve uzanma aktivitelerinde sınıflandırma %100 başarılıdır. Tırmanma aktivitesinde başarı %96 iken model aktivitenin %2,7' si yürüme, %0,91' i inme olarak sınıflandırmıştır. Oturma aktivitesinde başarı %76' dır. Oturma aktivitesinin %24' ü bekleme olarak, bekleme aktivitesinin %2,3' ü oturma, %0,76' sı tırmanma aktivitesi olarak tanımlanmış ve bekleme aktivitesinde başarı oranı %97' dir.

%20 test verisi ile yapılan uygulamada, inme ve uzanma aktivitelerinde başarı oranı %100' dür. Yürüme aktivitesinin %12' si inme olarak sınıflandırılırken başarı oranı %88' dur. Tırmanma aktivitesinin başarı oranı %66 iken %8,1' i yürüme, %26' sı inme aktivitesi olarak tahmin edilmiştir. Oturma aktivitesinin %1,2' si tırmanma, %0,4' ü uzanma, %1,6' sı bekleme, bekleme aktivitesinin %63' ü oturma olarak sınıflandırılmıştır. Oturma aktivitesinde başarı %97, bekleme aktivitesinde ise %37' dir.

Veri setinin %30' u ile elde edilen hata matrisinde, uzanma aktivitesinde %100 sınıflandırma başarısı elde edilmiştir. Tırmanma aktivitesinin %83' ü doğru sınıflandırılırken, aktivitenin %2,1' i yürüme, %15' ü inme olarak tanımlanmıştır. İnme aktivitesinde sınıflandırma başarısı %99, oturma aktivitesinde %98 ve bekleme aktivitesinde %51' dir. İnme aktivitesinin %0,88' i tırmanma olarak belirlenmiştir. Oturma aktivitesinin %0,96' sı bekleme, bekleme aktivitesinin %49' u oturma aktivitesi olarak sınıflandırılmıştır.

Çizelge 4.3. SVM ile elde edilen hata matrisleri örneklem gösterimi

	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma
	%90 Test %10 Eğitim					
Yürüme	145	0	0	0	0	0
Tırmanma	3	106	1	0	0	0
İnme	0	0	102	0	0	0
Oturma	0	0	0	92	29	0
Bekleme	0	1	0	3	127	0
Uzanma	0	0	0	0	0	126
	%80 Eğitim %20 Test					
Yürüme	238	0	32	0	0	0
Tırmanma	18	147	58	0	0	0
İnme	0	0	197	0	0	0
Oturma	0	3	0	243	4	1
Bekleme	0	0	0	164	95	0
Uzanma	0	0	0	0	0	270
	%70 Eğitim %30 Test					
Yürüme	360	0	54	0	0	0
Tırmanma	8	321	58	0	0	0
İnme	0	3	339	0	0	0
Oturma	0	3	0	409	4	1
Bekleme	0	0	0	227	235	0
Uzanma	0	0	0	0	0	428

Çizelge 4.3'te gösterilen hata matrisinde SVM modeli %10 test verisi ile yapılan uygulamada 145 yürüme aktivitesinin, 126 uzanma aktivitesinin ve 102 inme aktivitesinin tamamını doğru tahmin etmiştir. 110 tırmanma aktivitesinin 1'i inme, 3' ü yürüme olarak tahmin edilmiştir. 121 oturma aktivitesinin 29' u bekleme, aktivitesi olarak tahmin edilmiştir. 131 bekleme aktivitesinin ise 3' ü oturma aktivitesi ve 1' i tırmanma olarak belirlenmiştir.

Yapılan ikinci uygulamada %20 test verisi ile elde edilen sonuçlara göre model, 197 inme ve 270 uzanma aktivitelerinin tamamını doğru tahmin etmiştir. 270 yürüme aktivitesinin 32' si inme aktivitesi olarak tanımlanmıştır. 223 tırmanma aktivitesinin 18' i yürüme 58' i inme, 251 oturma aktivitesinin 4' ü bekleme, 3' ü tırmanma ve 1' i uzanma, 259 bekleme aktivitesinin 164' ü oturma aktivitesi olarak tahmin edilmiştir.

Veri setinin %30' u ile model test edildiğinde, 414 yürüme aktivitesinin 54' ü inme, 387 tırmanma aktivitesinin 8' i yürüme, 58' i inme, 342 inme aktivitesinin 3' ü tırmanma, 417 oturma aktivitesinin, 3' ü tırmanma, 4' ü bekleme, 462 bekleme aktivitesinin 227' si oturma aktivitesi olarak sınıflandırılmıştır. 428 uzanma aktivitesinin tamamı doğru tahmin edilmiştir.

En büyük sapma KNN modelinde oturma ve bekleme aktiviteleri arasında iken, SVM modelinde bekleme aktivitesinin oturma aktivitesiyle karıştırıldığı görülmektedir.

4.2.2 Test verisi ile metrikler

SVM modelinde elde edilen precision, recall, f1-score ve accuracy metrik değerleri ile test verisine ait örneklem sayıları Çizelge 4.4'te, gösterilmiştir.

Çizelge 4.4. SVM modeli ile elde edilen metrikler

	Precision %	Recall %	F1-Score %	Acc %	Örneklem Sayısı
	%90 Eğitim %10 Test				
Yürüme	0,98	1,00	0,99	0,95	145
Tırmanma	0,99	0,96	0,98		110
İnme	0,99	1,00	1,00		102
Oturma	0,97	0,76	0,85		121
Bekleme	0,81	0,97	0,89		131
Uzanma	1,00	1,00	1,00		126
	%80 Eğitim %20 Test				
Yürüme	0,93	0,88	0,90	0,81	270
Tırmanma	0,98	0,66	0,79		223
İnme	0,69	1,00	0,81		197
Oturma	0,60	0,97	0,74		251
Bekleme	0,96	0,37	0,53		259
Uzanma	1,00	1,00	1,00		270
	%70 Eğitim %30 Test				
Yürüme	0,98	0,87	0,92	0,85	414
Tırmanma	0,98	0,83	0,90		387
İnme	0,75	0,99	0,85		342
Oturma	0,64	0,98	0,78		417
Bekleme	0,98	0,51	0,67		462
Uzanma	1,00	1,00	1,00		428

Precision (hassaslık) metriğinde %10 test verisi ile en yüksek hassaslık değeri uzanma aktivitesinde gerçekleşmiştir. Uzanma aktivitesinin hassaslık değeri %100' dür. En düşük hassaslık değeri %81 ile bekleme aktivitesinde gerçekleşmiştir. Yürüme aktivitesinde %98, tırmanma aktivitesinde %99 ve oturma aktivitesinde %97 hassaslık değerleri elde edilmiştir.

Veri setinin %20' si ile test edildiğinde, Uzanma aktivitesinde %100 değeri elde edilmiştir. Tırmanma aktivitesinde %98, inme aktivitesinde %69 hassaslık değeri elde edilirken oturma aktivitesinde %60, bekleme aktivitesinde %96 değerleri elde edilmiştir.

Test verisi %30' a çıkarıldığında uzanma aktivitesinde hassaslık metriğinde başarı oranı %100' dür. Yürüme, tırmanma ve bekleme aktivitesinde %98 değeri elde edilirken, oturma aktivitesinde %64, inme aktivitesinde %75 değeri görülmüştür.

Aktivite sınıfları için yapılan tahminlerde doğruluk oranını gösteren recall metriğinde %10 test verisi ile en yüksek doğruluk oranları %100 ile yürüme, inme ve uzanma aktivitesinde gerçekleşmiştir. En düşük doğruluk oranı %76 ile oturma aktivitesinde gerçekleşmiştir. Tırmanma aktivitesinde %96, bekleme aktivitesinde %97 doğruluk oranları elde edilmiştir.

Veri setinin %20' si test için kullanıldığında recall metriği değerlerine göre, inme ve uzanma aktivitelerinin tamamının doğru tahmin edildiği görülmektedir. Yürüme aktivitesinde %88, tırmanma aktivitesinde %66 sınıflandırma başarısı görülmüştür. Oturma aktivitesinde %97 değeri elde edilirken en düşük değer %37 ile bekleme aktivitesinde elde edilmiştir.

Çizelge 4.4'te %30 test verisi ile elde edilen recall metriği değerleri gösterilmiştir. Uzanma aktivitesinde sınıflandırma %100 başarılıdır. İnme aktivitesinde %99, oturma aktivitesinde %98, yürüme aktivitesinde %87, tırmanma aktivitesinde %83 sınıflandırma başarısı elde edilmiştir. En düşük sınıflandırma başarısı %51 ile bekleme aktivitesinde görülmektedir. Precision ve recall metriklerinin bir ortalaması olan f1-score metriğinde %10 test verisi ile en yüksek değerler %100 ile inme ve uzanma aktivitelerinde görülmektedir. Yürüme aktivitesinde %99, tırmanma aktivitesinde %98, bekleme aktivitesinde %89 ve oturma aktivitesinde %85 değerleri elde edilmiştir.

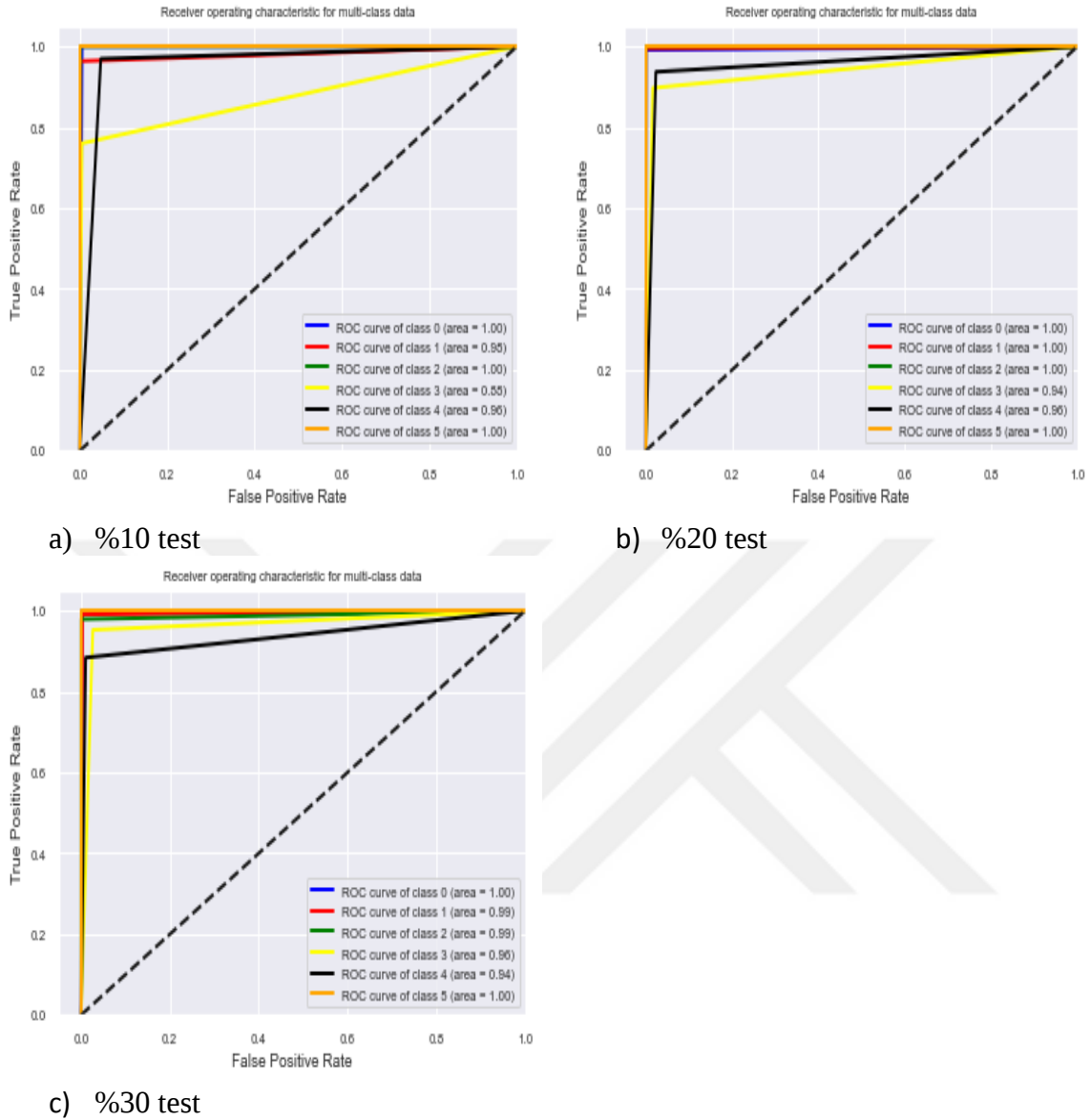
Veri setinin %20' si test için ayrıldığında uzanma aktivitesinde ortalama değer %100' dür. Yürüme aktivitesinde %90 ortalama elde edilirken, oturma aktivitesinde ortalama %74, tırmanma aktivitesinde %79, inme aktivitesinde %81 bekleme aktivitesinde %53 ortalama görülmüştür.

Yapılan üçüncü uygulamada test verisi oranı %30 olarak belirlenmiş ve f1-score değerleri uzanma aktivitesinde %100, yürüme aktivitesinde %92, tırmanma aktivitesinde %90, inme aktivitesinde %85 oranıyla elde edilirken en düşük değer %67 ile bekleme aktivitesinde görülmektedir.

SVM algoritması ile oluşturulan modelin ortalama doğruluk oranı (accuracy) %10 test verisi ile %95 iken %20 test verisi ile %81 ve %30 test verisi ile %85' dir. Sırasıyla 735, 1470 ve 2450 aktiviteden oluşan test verisinde sınıfların dağılımı örneklem sayısında gösterilmiştir.

4.2.3 Test verisi ile roc eğrileri

Veri setinde farklı test verileri ile SVM modeli kullanılarak elde edilen ve uygunluk oranlarını gösteren ROC eğrileri Şekil 4.4'te gösterilmiştir.



Şekil 4.4. SVM modeli ile elde edilen roc eğrileri a)%10 test b) %20 test c)%30 test

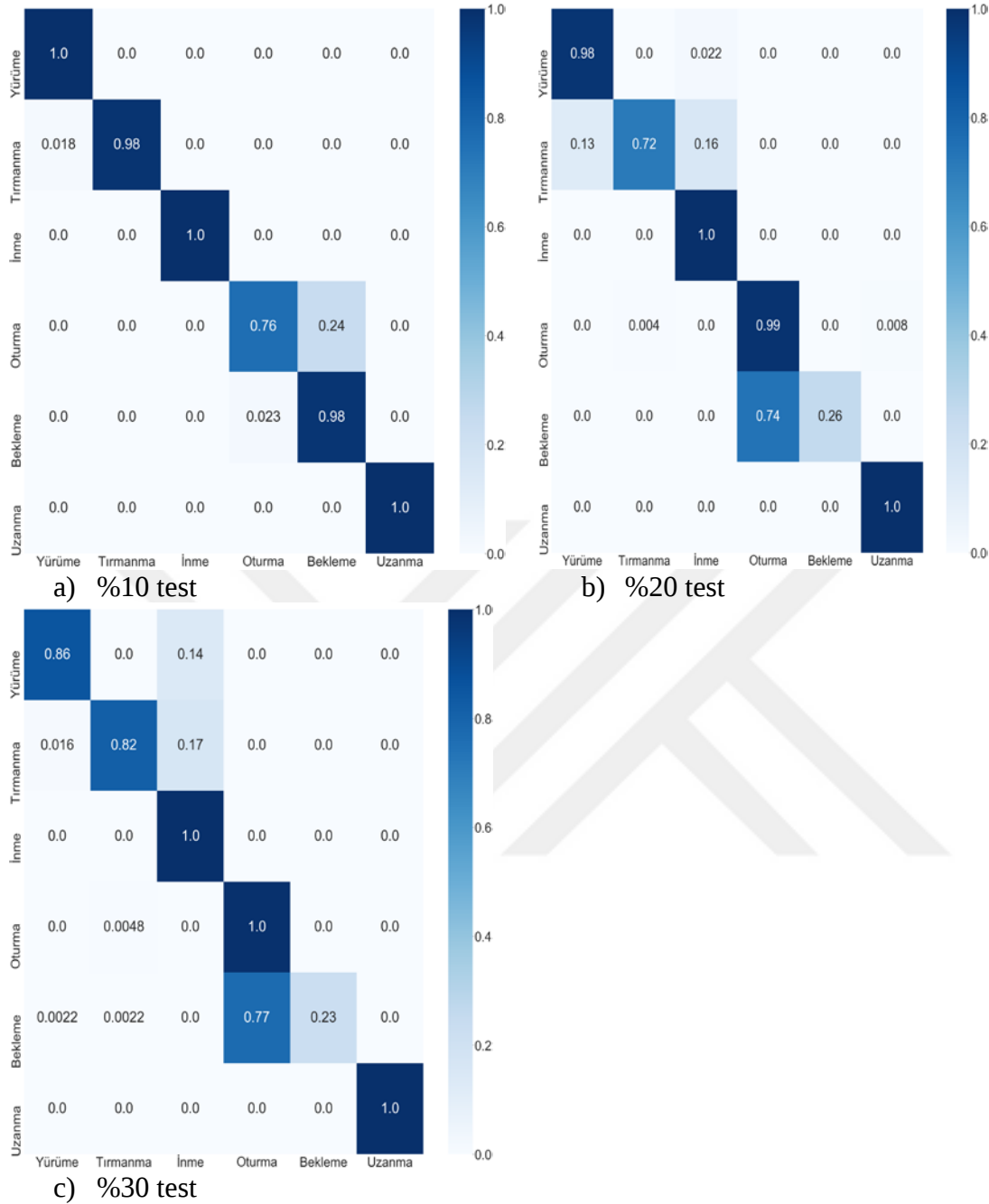
Şekil 4.4'e göre test verilerinin kabul edilebilirlik oranları, üç farklı test verisinde yürüme ve uzanma aktivitelerinde 1.00 olarak elde edilmiştir. Tırmanma ve inme aktivitelerinde 0.99-1.00 değerleri görülürken, oturma ve bekleme aktivitelerinde diğer aktivitelere göre daha düşük değerler elde edilmiştir. Sonuçlara göre test verilerinin makul düzeyde olduğu söylenebilir.

4.3 MLP Sınıflandırıcısı ile Elde Edilen Sonuçlar

Çok katmanlı algılayıcı algoritmasıyla yapılan uygulamalarda farklı katman ve nöron sayıları ile uygulamalar yapıp model optimize edilmiştir. Hata matrisleri ile yüzdelik gösterim yapılmış ve modelde elde edilen metriklerle roc eğrileri gösterilmiştir.

4.3.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi

MLP modelinde, alpha parametresi 1e-005, hidden_layer_size parametresi 20,10 olarak ayarlanmıştır. Bu durumda birinci katmanda 20 nöron, ikinci katmanda ise 10 nöron bulunmaktadır. Optimizasyon için 'sdg' ve aktivasyon fonksiyonu relu seçilmiştir. Modelin eğitimi için kullanılan iterasyon sayısı 1000 'dir. Elde edilen hata matrislerine ait yüzdelik oranları Şekil 4.5'te gösterilmiştir.



Şekil 4.5. MLP algoritması hata matrisleri a)%10 test b)%20 test c)%30 test

Matrise göre %10 test verisi ile yürüme, inme ve uzanma aktivitelerinde sınıflandırma %100 başarılıdır. Tırmanma aktivitesinde başarı %98 iken model aktivitenin %1,8' ini yürüme olarak aktivitesi olarak sınıflandırmıştır. Oturma aktivitesinin %0,24' ü bekleme olarak tanımlanırken %76 başarı oranı elde edilmiştir. Bekleme aktivitesinde başarı %98' dir. Bekleme aktivitesinin %2,3' ü ise oturma aktivitesi olarak sınıflandırılmıştır.

Şekil 4.5'e göre %20 test verisi ile inme ve uzanma aktivitelerinin sınıflandırılmasında başarı %100' dür. Yürüme aktivitesinde %98 sınıflandırma başarısı gözlenirken aktivitenin %2,2' si inme aktivitesi olarak belirlenmiştir. Tırmanma aktivitesinin %13' ü yürüme,

%16' sı inme aktivitesi olarak tahmin edilirken başarı oranı %72' dir. Oturma aktivitesinin %0,4' ü tırmanma, %0,8' i uzanma olarak sınıflandırılmış ve başarı oranı %99' dur. Bekleme aktivitesinin %74' ü oturma aktivitesi olarak tahmin edilirken başarı oranı %26' da kalmıştır.

Veri setinin %30' u test için ayrıldığında inme ve uzanma aktivitelerinde sınıflandırma başarısı %100' dür. Yürüme aktivitesinde %86 sınıflandırma başarısı gözlenirken aktivitenin %14' ü inme aktivitesi olarak sınıflandırılmıştır. Tırmanma aktivitesinde %82 sınıflandırma başarısı görülürken aktivitenin %17' si inme ve %1,6' sı yürüme aktivitesi olarak tahmin edilmiştir. Oturma aktivitesinin %0,48' i tırmanma aktivitesi olarak tahmin edilmiş ve başarı oranı %100' e yaklaşmıştır. Bekleme aktivitesinde %23' lük sınıflandırma başarısı görülürken %77' si oturma aktivitesi olarak tahmin edilmiştir. Aktivitenin %0,22' si yürüme, %0,22' si tırmanma olarak belirlenmiştir.

Çizelge 4.5. MLP ile elde edilen hata matrisleri örneklem gösterimi

	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma
	%90 Test %10 Eğitim					
Yürüme	145	0	0	0	0	0
Tırmanma	2	108	0	0	0	0
İnme	0	0	102	0	0	0
Oturma	0	0	0	92	29	0
Bekleme	0	0	0	3	128	0
Uzanma	0	0	0	0	0	126
	%80 Eğitim %20 Test					
Yürüme	264	0	6	0	0	0
Tırmanma	28	160	35	0	0	0
İnme	0	0	197	0	0	0
Oturma	0	1	0	248	0	2
Bekleme	0	0	0	192	67	0
Uzanma	0	0	0	0	0	270
	%70 Eğitim %30 Test					
Yürüme	355	0	59	0	0	0
Tırmanma	6	317	64	0	0	0
İnme	0	0	342	0	0	0
Oturma	0	2	0	415	0	0
Bekleme	1	1	0	356	104	0
Uzanma	0	0	0	0	0	428

Çizelge 4.5'te gösterilen hata matrisinde %10 test verisi ile MLP modeli 145 yürüme aktivitesinin, 102 inme aktivitesinin, 126 uzanma aktivitesinin tamamını doğru tahmin

etmiştir. 110 tırmanma aktivitesinin 2' si yürüme olarak tahmin edilmiştir. 121 oturma aktivitesinin 29' u bekleme aktivitesi olarak sınıflandırılırken, 131 bekleme aktivitesinin 3' ü oturma aktivitesi olarak tahmin edilmiştir.

Yapılan ikinci uygulamada %20 test verisi kullanılmıştır. Buna göre 197 inme aktivitesi ve 270 uzanma aktivitesinin tamamı doğru tahmin edilmiştir. 270 yürüme aktivitesinin 6' sı inme, 223 tırmanma aktivitesinin 28' i yürüme 35' i inme, 251 oturma aktivitesinin 1' si tırmanma ve 2' si uzanma, 259 bekleme aktivitesinin 192' si oturma aktivitesi olarak belirlenmiştir.

Çizelge 4.5'e göre %30 test verisi ile yapılan uygulamada, 342 inme aktivitesinin ve 428 uzanma aktivitesinin tamamı doğru sınıflandırılmıştır. 414 yürüme aktivitesinin 59' u inme, 387 tırmanma aktivitesinin 6'sı yürüme 64' ü inme, 417 oturma aktivitesinin 2' si tırmanma, 362 bekleme aktivitesinin 356'sı oturma olarak sınıflandırılmıştır.

Modelden elde edilen sonuçlarda en büyük sapmalar bekleme aktivitesinde görülmektedir. Yapılan uygulamalarda %20 ve %30' luk test verilerinde elde edilen sonuçlara göre bekleme aktivitesinin çoğunlukla oturma aktivitesi olarak sınıflandırıldığı görülmektedir.

4.3.2 Test verisi ile metrikler

MLP modeli ile elde edilen precision, recall, f1-score ve accuracy metrik değerleri ile test verilerine ait örneklem sayıları Çizelge 4.6' da gösterilmiştir.

Çizelge 4.6. MLP modeli ile elde edilen metrikler

	Precision %	Recall %	F1-Score %	Acc %	Örneklem Sayısı
	%90 Eğitim %10 Test				
Yürüme	0,99	1,00	0,99	0,95	145
Tırmanma	1,00	0,98	0,99		110
İnme	1,00	1,00	1,00		102
Oturma	0,97	0,76	0,85		121
Bekleme	0,82	0,98	0,89		131
Uzanma	1,00	1,00	1,00		126
	%80 Eğitim %20 Test				
Yürüme	0,90	0,98	0,94	0,82	270
Tırmanma	0,99	0,72	0,83		223
İnme	0,83	1,00	0,91		197
Oturma	0,56	0,99	0,72		251
Bekleme	1,00	0,26	0,41		259
Uzanma	0,99	1,00	1,00		270
	%70 Eğitim %30 Test				
Yürüme	0,98	0,86	0,91	0,80	414
Tırmanma	0,99	0,82	0,90		387
İnme	0,74	1,00	0,85		342
Oturma	0,54	1,00	0,70		417
Bekleme	1,00	0,23	0,37		462
Uzanma	1,00	1,00	1,00		428

Veri setinin %10' u ile yapılan test uygulamasında en yüksek hassaslık değeri tırmanma, inme ve uzanma aktivitelerinde gerçekleşmiştir. Tırmanma, inme ve uzanma aktivitelerinin hassaslık değeri %100' dür. En düşük hassaslık değeri %82 ile bekleme aktivitesinde gerçekleşmiştir. Yürüme aktivitesinde %99, oturma aktivitesinde %97 hassaslık değerleri elde edilmiştir.

Yapılan ikinci uygulamada %20 test verisi ile çalışılmıştır. Çizelge 9' a göre, en yüksek hassaslık değeri %100 ile bekleme aktivitesinde görülmüştür. Yürüme aktivitesinde %90, tırmanma ve uzanma aktivitelerinde %99, inme aktivitesinde %83 ve oturma aktivitesinde %56 ile en düşük hassaslık değeri görülmektedir.

Hassaslık metriğinde %30 test verisi ile elde edilen sonuçlara göre, bekleme ve uzanma verilerinde görülen değerler %100' dür. Yürüme %98, tırmanma %99, inme aktivitesinde %74 hassaslık değeri görülürken oturma aktivitesinde %54 ile en düşük değer görülmüştür. Özgüllük (recall) metriğinde sınıflara göre dağılım görülmektedir. Veri setinde %10 test verileri ile en yüksek doğruluk oranları %100 ile yürüme, inme ve uzanma aktivitelerinde gerçekleşmiştir. En düşük oran %76 ile oturma aktivitesinde gerçekleşmiştir. Tırmanma ve bekleme aktivitelerinde %98 oranları elde edilmiştir.

İkinci uygulamada verilerin %20' si test verisi olarak kullanılmıştır. İnme ve uzanma aktivitelerinde %100, oturma aktivitesinde %99, yürüme aktivitesinde %98, tırmanma aktivitesinde %72 oranları elde edilirken en düşük değer %26 ile bekleme aktivitesinde elde edilmiştir.

Çizelge 4.6'da %30 test oranı ile elde edilen özgüllük metriğinde %100 değeri inme, oturma ve uzanma aktivitelerinde görülmüştür. Yürüme aktivitesinde %86, tırmanma aktivitesinde %82 oranları görülürken en düşük değer %23 ile bekleme aktivitesinde görülmektedir.

Precision ve recall metriklerinin ortalamasını ifade eden f1-score metriğinde %10 test verisinde en yüksek değer %100 ile inme ve uzanma aktivitelerinde görülmektedir. Yürüme ve tırmanma aktivitelerinde %99, bekleme aktivitesinde %89 ve oturma aktivitesinde %85 değerleri elde edilmiştir.

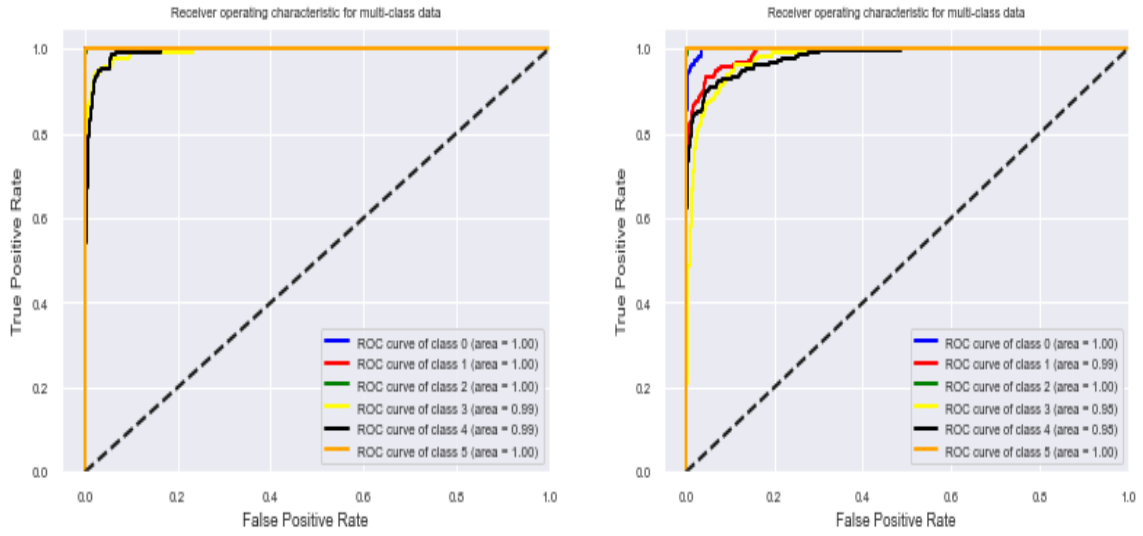
Yapılan ikinci uygulamada %20 test verisi ile elde edilen %100' lük ortalama uzanma aktivitesinde görülmektedir. En düşük ortalama %41 ile bekleme aktivitesinde görülmüştür. Yürüme aktivitesinde %94, inme %91, tırmanma %83 ve oturma aktivitesinde ortalama %72' dir.

Üçüncü uygulamada %30' luk test verisiyle, en yüksek ortalama %100 ile uzanma aktivitesinde görülmektedir. En düşük ortalama %37 ile bekleme aktivitesinde görülmektedir. Yürüme %91, tırmanma %90, inme %85, oturma aktivitesinde ise ortalama %70' tir.

MLP algoritması ile oluşturulan modelin ortalama doğruluk oranları (accuracy) %95, %82 ve %80 olarak bulunmuştur. 735, 1470 ve 2450 aktiviteden oluşan test verilerinde sınıfların dağılımı örneklem sayısında gösterilmiştir.

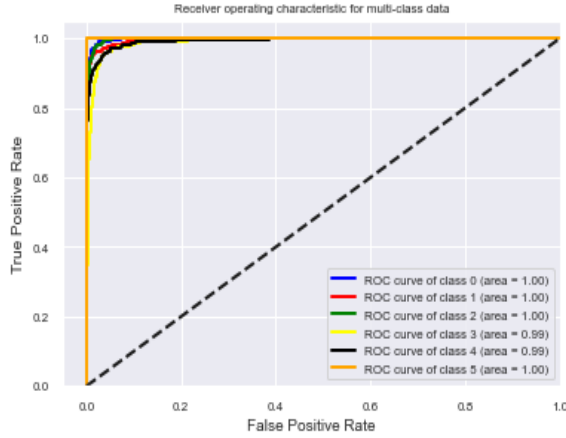
4.3.3 Test verisi ile roc eğrisi

Test verilerinin MLP modeli ile elde edilen ve doğruluk oranlarını gösteren ROC eğrileri Şekil 4.6'da gösterilmiştir.



a) %10 test

b) %20 test



c) %30 test

Şekil 4.6. MLP modeli ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test

Şekil 4.6'ya göre test verilerinin doğruluk oranları, tüm aktivitelerde 1.0' e yakın olarak elde edilmiştir. Oturma ve bekleme aktivitelerinde %20 test verisi ile %95 değeri görülürken diğer uygulamalarda %99 olduğu görülmektedir. Sonuçlara göre test verilerinin kabul edilebilir olduğu söylenebilir.

4.4 RF Sınıflandırıcısı ile Elde Edilen Sonuçlar ve ROC Eğrisi

Rastgele orman algoritmasıyla yapılan uygulamalarda parametre seçimleri denenmiş en iyi sonuçlar elde edilene kadar model optimize edilmiştir. Hata matrisleri ile yüzdelik gösterim yapılmış ve modelde elde edilen metriklerle roc eğrileri gösterilmiştir.

4.4.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi

RF modelinde, criterion parametresi “gini” seçilmiştir. Elde edilen hata matrisleri yüzdelik oranları Şekil 4.7’de gösterilmiştir.



Şekil 4.7. RF algoritması hata matrisleri a)%10 test b)%20 test c)%30 test

Şekil 4.7’deki matrislere göre, %10 test verisi ile yürüme ve uzanma aktivitelerinde sınıflandırma %100 başarılıdır. İnme aktivitesinde başarı %98 iken aktivitenin %2’si yürüme aktivitesi olarak tahmin edilmiştir. Tırmanma aktivitesinin %25’ i yürüme ve %20’si inme aktivitesi olarak sınıflandırılırken başarı oranı %55’ tir. Oturma aktivitesinin

%2,5'i yürüme, %24' ü bekleme, bekleme aktivitesinin %11' i oturma ve %1,5' i yürüme aktivitesi olarak tahmin edilmiştir. Oturma aktivitesinde sınıflandırma başarıları %74 iken bekleme aktivitesinde %88 olduğu görülmektedir.

Test verisi oranı %20 olarak belirlendiğinde, uzanma aktivitesinde başarı %100' dür. İnme ve oturma aktivitelerinde %99' luk sınıflandırma başarıları görülürken, inme aktivitesinin %0,51' i yürüme, oturma aktivitesinin %0,8' i yürüme aktivitesi olarak tahmin edilmiştir. Yürüme aktivitesinin başarı oranı %77' dir ve aktivitenin %23' ü inme aktivitesi olarak sınıflandırılmıştır. Tırmanma aktivitesinin %29' u yürüme, %31' i inme olarak sınıflandırılmış ve başarı oranı %40' tır. Bekleme aktivitesinin %77' si oturma aktivitesi olarak tahmin edilmiş ve başarı oranı %19' dur. Bekleme aktivitesinin %3,9' u yürüme aktivitesi olarak tahmin edilmiştir.

Şekil 4.7'ye göre test verilerinin %30' u ile yapılan uygulamada uzanma aktivitesinde sınıflandırma başarıları %100' dür. Yürüme aktivitesinin %19' u inme, %0,24' ü tırmanma olarak sınıflandırılırken başarı oranı %81'dir. Tırmanma aktivitesinde başarı %55 iken aktivitenin %26' sı yürüme ve %19' u inme aktivitesi olarak tahmin edilmiştir. İnme aktivitesinde başarı oranı %95' tir. Aktivitenin %2,9' u tırmanma, %1,8' i yürüme aktivitesi olarak tahmin edilmiştir. Oturma aktivitesinin %0,72' si yürüme aktivitesi olarak sınıflandırılmıştır ve başarı oranı %99' dur. Bekleme aktivitesinin %78' i oturma aktivitesi, ve %6,9 ' u yürüme aktivitesi olarak belirlenirken başarı oranı %15' te kalmıştır.

Çizelge 4.7. RF ile elde edilen hata matrisleri örneklem gösterimi

	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma
	%90 Test %10 Eğitim					
Yürüme	145	0	0	0	0	0
Tırmanma	27	61	22	0	0	0
İnme	2	0	100	0	0	0
Oturma	3	0	0	89	29	0
Bekleme	2	0	0	14	115	0
Uzanma	0	0	0	0	0	126
	%80 Eğitim %20 Test					
Yürüme	207	0	63	0	0	0
Tırmanma	65	89	69	0	0	0
İnme	1	0	196	0	0	0
Oturma	2	0	0	249	0	0
Bekleme	10	0	0	199	50	0
Uzanma	0	0	0	0	0	270
	%70 Eğitim %30 Test					
Yürüme	335	1	78	0	0	0
Tırmanma	101	212	74	0	0	0
İnme	6	10	326	0	0	0
Oturma	3	0	0	414	0	0
Bekleme	32	0	0	362	68	0
Uzanma	0	0	0	0	0	428

Çizelge 4.7’de gösterilen hata matrislerinde %10 test verisi ile RF modeli 126 uzanma aktivitesinin ve 145 yürüme aktivitesinin tamamını doğru tahmin etmiştir. 110 tırmanma aktivitesinin 27’ si yürüme ve 22’ si inme, 102 inme aktivitesinin 2’ si yürüme, 121 oturma aktivitesinin 29’ u bekleme, 131 bekleme aktivitesinin 14’ ü oturma ve 2’ si yürüme olarak tahmin edilmiştir.

Test verisi oranı %20 alındığında, 270 uzanma aktivitesinin tamamı doğru tahmin edilmiştir. 251 oturma aktivitesinin 2’ si yürüme, 259 bekleme aktivitesinin 199’ u oturma ve 10’ u yürüme aktivitesi olarak tanımlanmıştır. 223 tırmanma aktivitesinin 65’ i yürüme ve 69’ u inme olarak tahmin edilirken 197 inme aktivitesinin 1’ i yürüme aktivitesi olarak belirlenmiştir.

Çizelge 4.7’ye göre test verisi oranı %30’ a çıkarıldığında uzanma 428 uzanma aktivitesinin tamamı doğru tahmin edilmiştir. 414 yürüme aktivitesinin 1’ i tırmanma ve 78’ i inme aktivitesi, 387 tırmanma aktivitesinin 101’ i yürüme, 74’ ü inme, 342 inme

aktivitesinin 6' s' yürüme ve 10' u tırmanma, 417 oturma aktivitesinin 3' ü yürüme, 462 bekleme aktivitesinin 362' si oturma ve 32' si yürüme aktivitesi olarak sınıflandırılmıştır. KNN, SVM ve MLP modellerinde olduğu gibi RF modelindeki sonuçlarda da oturma ve bekleme verilerinin modeller tarafından karıştırıldığı görülmektedir. En büyük sapmalar bu iki sınıfta gerçekleşmiştir.

4.4.2 Test verisi ile metrikler

RF modeli ile elde edilen precision, recall, f1-score ve accuracy metrik değerleri ile test verilerine ait örneklem sayıları Çizelge 4.8'de gösterilmiştir.

Çizelge 4.8. RF modeli ile elde edilen metrikler

	Precision %	Recall %	F1-Score %	Acc %	Örneklem Sayısı
	%90 Eğitim %10 Test				
Yürüme	0,81	1,00	0,90	0,87	145
Tırmanma	1,00	0,55	0,71		110
İnme	0,82	0,98	0,89		102
Oturma	0,86	0,74	0,79		121
Bekleme	0,80	0,88	0,84		131
Uzanma	1,00	1,00	1,00		126
	%80 Eğitim %20 Test				
Yürüme	0,73	0,77	0,75	0,72	270
Tırmanma	1,00	0,40	0,57		223
İnme	0,60	0,99	0,75		197
Oturma	0,56	0,99	0,71		251
Bekleme	1,00	0,19	0,32		259
Uzanma	1,00	1,00	1,00		270
	%70 Eğitim %30 Test				
Yürüme	0,70	0,81	0,75	0,73	414
Tırmanma	0,95	0,55	0,70		387
İnme	0,68	0,95	0,80		342
Oturma	0,53	0,99	0,69		417
Bekleme	1,00	0,15	0,26		462
Uzanma	1,00	1,00	1,00		428

RF modeli ile %10 test verisinde Precision (hassaslık) metriğinde en yüksek hassaslık değeri tırmanma ve uzanma aktivitelerinde gerçekleşmiştir. Tırmanma ve uzanma aktivitelerinin hassaslık değeri %100' dür. En düşük hassaslık değeri %80 ile bekleme aktivitesinde gerçekleşmiştir. Yürüme aktivitesinde %81, inme aktivitesinde %82, oturma aktivitesinde %86 hassaslık değerleri elde edilmiştir.

Yapılan ikinci uygulamada test verileri %20 olarak alınmıştır. Buna göre bekleme ve uzanma aktivitelerinin hassaslık değeri %100' dür. En düşük hassaslık değeri %56 ile oturma aktivitesinde görülmüştür. Yürüme aktivitesinde %73, inme aktivitesinde %60 değeri elde edilmiştir.

Veri setinin %30' u ile yapılan uygulamada bekleme ve uzanma aktivitelerinde hassaslık değeri %100' dür. En düşük hassaslık değeri %53 ile oturma aktivitesinde görülmektedir. Yürüme aktivitesinde %70, tırmanma aktivitesinde %95, inme aktivitesinde %68 değerleri elde edilmiştir.

Sınıflara göre doğruluk oranlarını gösteren recall metriğinde %10 test verisi ile en yüksek doğruluk oranı %100 ile yürüme ve uzanma aktivitelerinde gerçekleşmiştir. En düşük doğruluk oranı %55 ile tırmanma aktivitesinde gerçekleşmiştir. İnme aktivitesinde %98, oturma aktivitesinde %74, bekleme aktivitesinde %88 sınıflandırma doğruluğu elde edilmiştir.

Test verileri %20 oranına yükseltildiğinde uzanma aktivitesinde sınıflandırma başarısının %100 olduğu görülmektedir. İnme ve oturma aktivitelerinde %99 başarı elde edilirken, en düşük başarı oranı %19 ile bekleme aktivitesindedir. Yürüme aktivitesinde %77, tırmanma aktivitesinde %40 sınıflandırma başarısı elde edilmiştir.

Yapılan üçüncü uygulamada test verisi oranı %30 olarak seçilmiştir. Uzanma aktivitesinde %100 sınıflandırma başarısı elde edilmiştir. Yürüme aktivitesinde %81, tırmanma aktivitesinde 55, inme aktivitesinde %95, oturma aktivitesinde %99 sınıflandırma başarısı gözlenmiştir. En düşük sınıflandırma başarısı %15 ile bekleme aktivitesinde görülmektedir.

F1-score metriğinde gözlemlenen ortalamalarda %10 test verisi ile en yüksek değer %100 ile uzanma aktivitesinde görülmektedir. En düşük ortalama değer tırmanma aktivitesinde %71'dir. Yürüme aktivitesinde %90, inme aktivitesinde %89, oturma aktivitesinde %79 ve bekleme aktivitelerinde %84 ortalama değer görülmektedir.

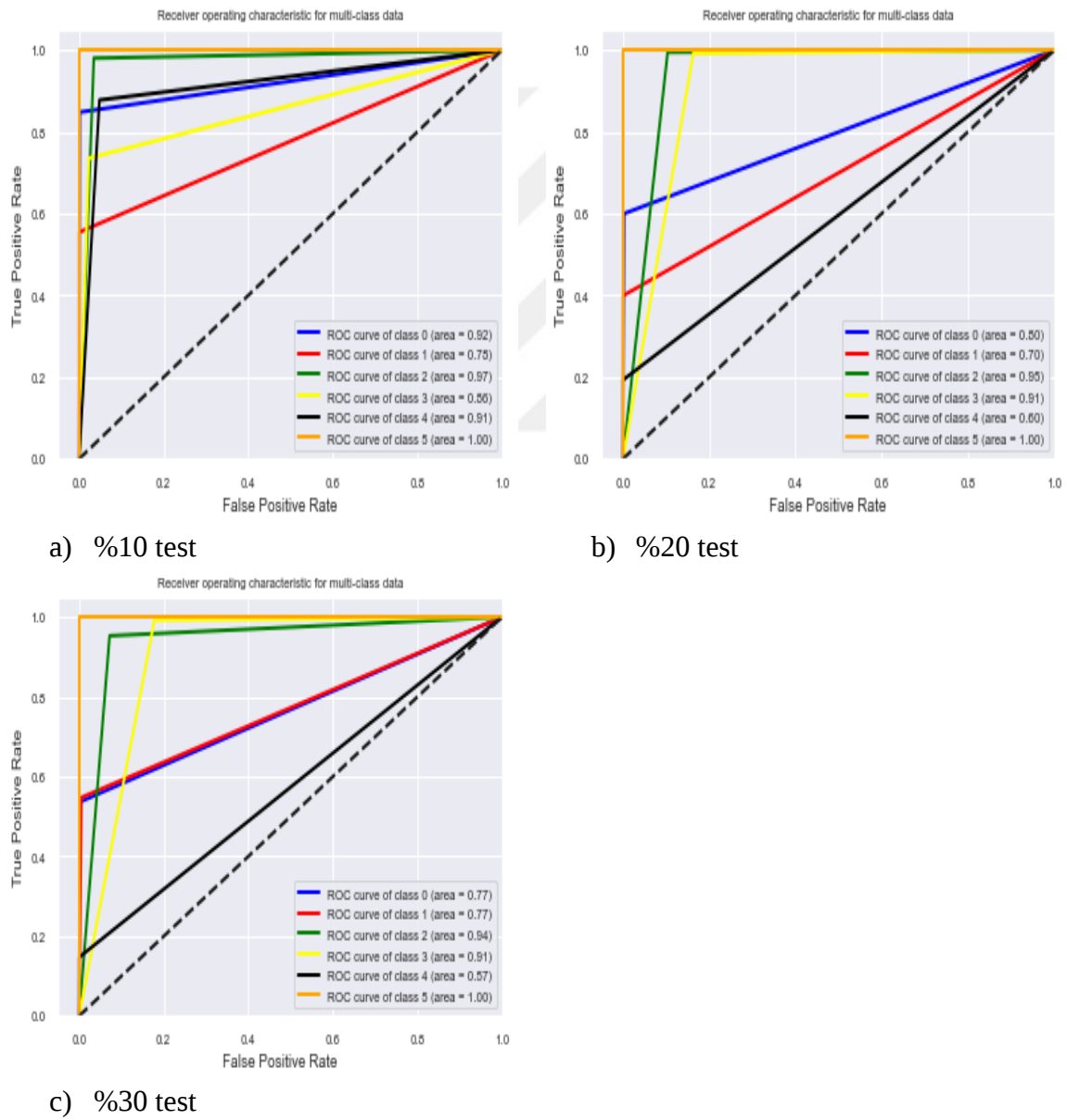
Test verisi %20 olarak seçildiğinde uzanma aktivitesinde %100, yürüme ve inme aktivitelerinde %75, oturma aktivitesinde %71 ortalama değerleri görülmüştür. En düşük ortalama değer %57 ile tırmanma aktivitesindedir.

Yapılan üçüncü uygulamada test verisi %30 olarak alınmıştır. Uygulamada en yüksek ortalama %100 ile uzanma aktivitesindedir. En düşük ortalama değer %26 ile bekleme aktivitesinde görülmüştür. Yürüme aktivitesinde %75, tırmanma aktivitesinde %70, inme aktivitesinde %80, oturma aktivitesinde %69 ortalama değerler görülmüştür.

RF algoritması ile oluşturulan modelin ortalama doğruluk oranı (accuracy), test verisi %10 seçildiğinde %87, %20 seçildiğinde %72 ve %30 seçildiğinde %73' tür. 735, 1470 ve 2450 aktiviteden oluşan test verilerinde sınıfların dağılımı örneklem sayısında gösterilmiştir.

4.4.3 Test verisi ile roc eğrileri

Test verilerinin RF modeli ile elde edilen ve doğruluk oranlarını gösteren ROC eğrileri Şekil 4.8'de gösterilmiştir.

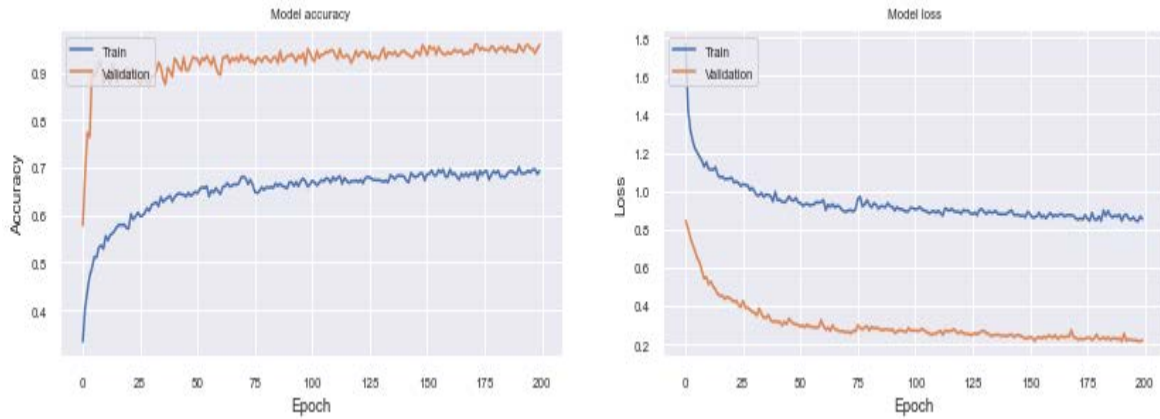


Şekil 4.8. RF modeli ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test

Şekil 4.8'e göre test verilerinin tutarlılığı, uzmanma aktivitesinde 1,0 olarak elde edilmiştir. Diğer aktivitelerde değerler 0,57-0,97 arasında değişmektedir. Test verileri %30 olarak alındığında bekleme aktivitesinde 0,57 ile referans köşegen çizgisine yaklaşıldığı görülmektedir. Sonuçlara göre test verilerinin kabul edilebilir olduğu söylenebilir.

4.5 Derin Öğrenme Modeli ile Elde Edilen Sonuçlar ve ROC Eğrisi

Model oluşturulurken tensorflow arka planını kullanan keras kütüphanesinden yararlanılmıştır. Modelde 4 katmanlı bir yapı kullanılırken ilk katmanda 16 nöron ve aktivasyon fonksiyonu olarak 'tanh', ikinci katmanda 9 nöron ve 'tanh' aktivasyon fonksiyonu, üçüncü katmanda 32 nöron ve 'tanh' aktivasyon fonksiyonu, dördüncü katmanda 6 nöron ve 'softmax' aktivasyon fonksiyonu kullanılmıştır. Başlangıçta öğrenme oranı 0,001 seçilmiştir. Batch_size parametresi 4, epoch sayısı 200 seçilmiş ve model optimize edilmeye çalışılmıştır. Model optimizasyonu için kütüphanede bulunan 'SGD' yöntemi, kayıp fonksiyonu olarak çoklu sınıflandırmayı destekleyen 'categorical_crossentropy' kullanılmıştır. Üç farklı uygulamada model eğitimi sırasında doğrulama(validation) için eğitim verilerinin %10' u ayrılmıştır. Model eğitimi sırasında 200 epoch ile elde edilen eğitim ve validasyon eğrileri Şekil 4.9'da gösterilmiştir.

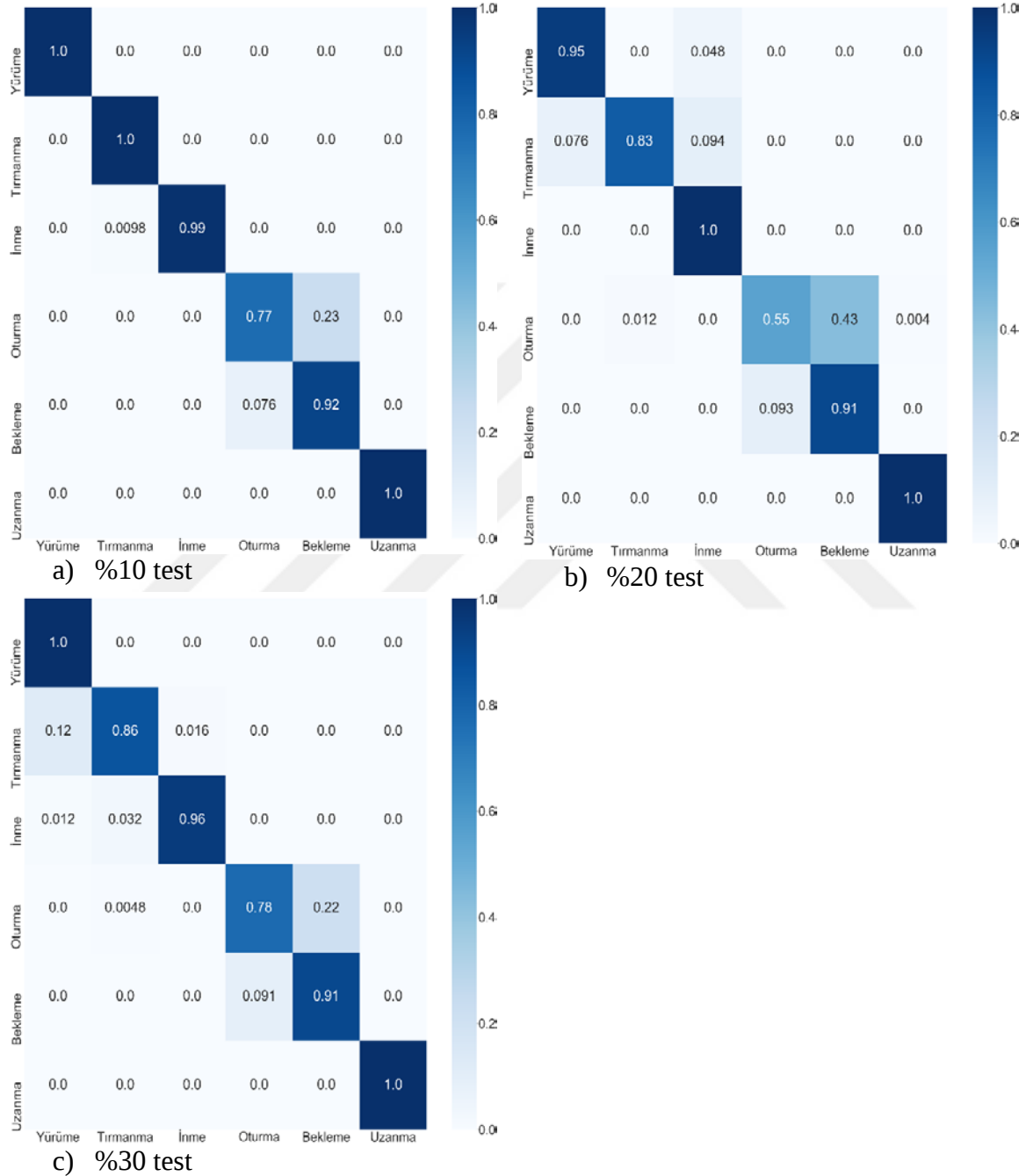


Şekil 4.9. Derin Öğrenme Modeli Eğitim ve Doğruluk Eğrileri

Şekil 4.9'da kayıp fonksiyonu eğrisi ve doğruluk oranı eğrileri incelendiğinde eğitim ve doğrulama sonuçlarında paralellik görülürken, her bir epoch sonunda eğitim ve doğruluğu artarken kayıp değeri azalmıştır.

4.5.1 Test verisi ile hata matrisleri yüzdelik ve örneklem gösterimi

Elde edilen hata matrislerine ait yüzdelik oranları Şekil 4.10'da gösterilmiştir. Derin öğrenme modelinde de üç farklı uygulama yapılmıştır.



Şekil 4.10. Derin öğrenme modeli hata matrisleri a)%10 test b)%20 test c)%30 test

Şekil 4.10'daki matrislere göre, %10 test verisi ile yapılan uygulamada yürüme, tırmanma ve uzanma aktivitelerinde sınıflandırma %100 başarılıdır. İnme aktivitesinde başarı %99 iken aktivitenin %0,98'i tırmanma aktivitesi olarak tahmin edilmiştir. Oturma aktivitesinin

%23'ü bekleme, bekleme aktivitesinin %7,6' sı oturma aktivitesi olarak tahmin edilmiştir. Oturma aktivitesinde sınıflandırma başarısı %77 iken bekleme aktivitesinde %92 olduğu görülmektedir.

Test verisi oranı %20 olarak belirlendiğinde, inme ve uzanma aktivitelerinde başarı %100' dür. Yürüme aktivitesinde %95, tırmanma aktivitesinde %83, oturma aktivitesinde %55, bekleme aktivitesinde %91 sınıflandırma başarısı elde edilmiştir. Yürüme aktivitesinin %4,8' i inme, tırmanma aktivitesinin %9,4' ü inme ve %7,6' sı yürüme aktivitesi olarak sınıflandırılmıştır. Oturma aktivitesinde sınıflandırma başarısı %55 iken bekleme aktivitesinde %91 olduğu görülmektedir. Oturma aktivitesinin %43' ü bekleme, %0,4' ü uzanma, %1,2' si tırmanma aktivitesi olarak sınıflandırılırken bekleme aktivitesinin %9,3' ü oturma aktivitesi olarak tahmin edilmiştir.

Şekil 4.10'a göre test verilerinin %30' u ile yapılan uygulamada yürüme ve uzanma aktivitelerinde sınıflandırma başarısı %100' dür. Tırmanma aktivitesinin %12' si yürüme, %1,6' sı inme aktivitesi olarak tahmin edilirken başarı oranı %86' dır. İnme aktivitesinin %1,2' si yürüme, %3,2' si tırmanma aktivitesi olarak belirlenmiş ve başarı oranı %96' dır. Oturma aktivitesinde %78 sınıflandırma başarısı elde edilirken aktivitenin, %22' si bekleme, %0,48' i tırmanma olarak tahmin edilmiştir. Bekleme aktivitesinin %9,1' i oturma aktivitesi olarak tahmin edilirken başarı oranının %91 olduğu görülmektedir.

Çizelge 4.9. Derin öğrenme modeli ile elde edilen hata matrisleri örneklem gösterimi

	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma
	%90 Test %10 Eğitim					
Yürüme	145	0	0	0	0	0
Tırmanma	0	110	0	0	0	0
İnme	0	1	101	0	0	0
Oturma	0	0	0	93	28	0
Bekleme	2	0	0	10	121	0
Uzanma	0	0	0	0	0	126
	%80 Eğitim %20 Test					
Yürüme	257	0	13	0	0	0
Tırmanma	17	185	21	0	0	0
İnme	0	0	197	0	0	0
Oturma	0	3	0	139	108	0
Bekleme	0	0	0	24	235	0
Uzanma	0	0	0	0	0	270
	%70 Eğitim %30 Test					
Yürüme	414	0	0	0	0	0
Tırmanma	47	334	6	0	0	0
İnme	4	11	327	0	0	0
Oturma	0	2	0	324	91	0
Bekleme	0	0	0	42	420	0
Uzanma	0	0	0	0	0	428

Çizelge 4.9’da gösterilen hata matrislerinde %10 test verisi ile derin öğrenme modeli 145 yürüme aktivitesinin, 110 tırmanma aktivitesinin ve 126 uzanma aktivitesinin tamamını doğru tahmin etmiştir. 102 inme aktivitesinin 1’ i tırmanma, 121 oturma aktivitesinin 28’ i bekleme, 131 bekleme aktivitesinin 10’ u oturma ve 2’ si yürüme olarak tahmin edilmiştir. Test verisi oranı %20 alındığında, 270 uzanma aktivitesi ve 197 inme aktivitesinin tamamı doğru tahmin edilmiştir. 223 tırmanma aktivitesinin 17’ si yürüme ve 21’ i inme aktivitesi olarak tahmin edilmiştir. 251 oturma aktivitesinin 3’ ü inme ve 108’ i bekleme, 259 bekleme aktivitesinin 24’ ü oturma aktivitesi olarak tanımlanmıştır. 270 yürüme aktivitesinin 13’ ü inme aktivitesi olarak tahmin edilmiştir.

Çizelge 4.9’a göre test verisi oranı %30’ a çıkarıldığında 414 yürüme aktivitesi ve 428 uzanma aktivitesinin tamamı doğru tahmin edilmiştir. 387 tırmanma aktivitesinin 47’ si yürüme, 6’ sı inme, 342 inme aktivitesinin 4’ ü yürüme ve 11’ i tırmanma, 417 oturma aktivitesinin 2’ si tırmanma ve 91’ i bekleme, 462 bekleme aktivitesinin 42’ si oturma aktivitesi olarak sınıflandırılmıştır.

Klasik makine öğrenmesi modellerinde olduğu gibi derin öğrenme modelindeki sonuçlarda da oturma ve bekleme verilerinin modeller tarafından karıştırıldığı görülmektedir.

4.5.2 Test verisi ile metrikler

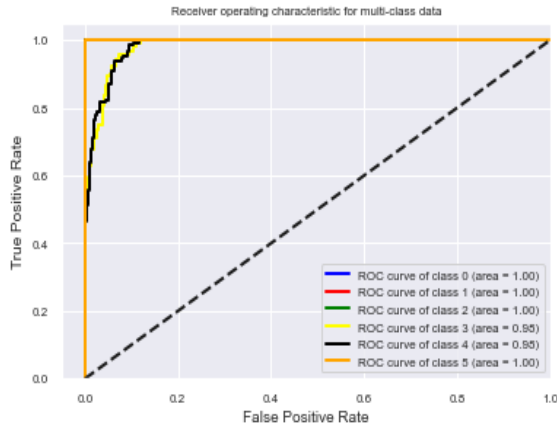
Derin öğrenme modeli ile elde edilen precision, recall, f1-score ve accuracy metrik değerleri ile test verilerine ait örneklem sayıları Çizelge 4.10’da gösterilmiştir.

Çizelge 4.10. Derin öğrenme modeli ile elde edilen metrikler

	Precision %	Recall %	F1-Score %	Acc %	Örneklem Sayısı
	%90 Eğitim %10 Test				
Yürüme	1,00	1,00	1,00	0,95	145
Tırmanma	0,99	1,00	1,00		110
İnme	1,00	0,99	1,00		102
Oturma	0,90	0,77	0,83		121
Bekleme	0,81	0,92	0,86		131
Uzanma	1,00	1,00	1,00		126
	%80 Eğitim %20 Test				
Yürüme	0,94	0,95	0,94	0,87	270
Tırmanma	0,98	0,83	0,90		223
İnme	0,85	1,00	0,92		197
Oturma	0,85	0,55	0,67		251
Bekleme	0,69	0,91	0,78		259
Uzanma	1,00	1,00	1,00		270
	%70 Eğitim %30 Test				
Yürüme	0,89	1,00	0,94	0,92	414
Tırmanma	0,96	0,86	0,91		387
İnme	0,98	0,96	0,97		342
Oturma	0,89	0,78	0,83		417
Bekleme	0,82	0,91	0,86		462
Uzanma	1,00	1,00	1,00		428

4.5.3 Test verisi ile roc eğrileri

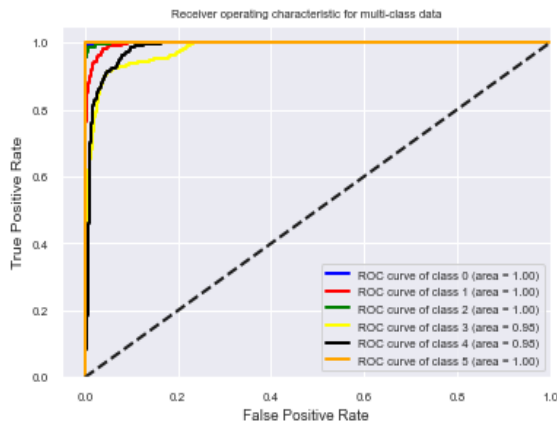
Test verilerinin derin öğrenme modeli ile elde edilen ve doğruluk oranlarını gösteren ROC eğrileri Şekil 4.11’de gösterilmiştir.



a) %10 test



b) %20 test



c) %30 test

Şekil 4.11. Derin öğrenme ile elde edilen roc eğrileri a)%10 test b)%20 test c)%30 test

Şekil 4.11'e göre test verilerinin tutarlılığı, yürüme, tırmanma, inme ve uzanma aktivitelerinde 1,0 olarak elde edilmiştir. Diğer aktivitelerde değerler 0,95-0,97 arasında değişmektedir. Sonuçlara göre test verilerinin kabul edilebilir olduğu söylenebilir.

5. SONUÇ VE ÖNERİLER

Bu çalışmada, akıllı telefonlarda bulunan ivmeölçer ve jiroskop sensör verileri ile makine öğrenmesi yöntemleri kullanarak güvenlik personellerinin fiziksel aktivitelerinin tanımlanması ortaya konmuştur. Güvenlik personellerinin zorlu şartlarda görev yaptığı düşünüldüğünde aktivitelerinin eş zamanlı takibinin yapılması personellerin güvenliği açısından önem arz etmektedir.

Çalışmada altı farklı aktiviteye ait ivmeölçer ve jiroskop sensör verileri içeren etiketlenmiş veri seti bölünerek farklı oranlarda eğitim ve test amaçlı kullanılmıştır. Özellik seçimi (feature selection) uygulanarak klasik makine öğrenmesi algoritmalarının yanı sıra bir derin öğrenme modeli kullanılmıştır. Modeller eğitildikten sonra daha önce hiç görmediği verilerle test edilmiştir.

Test sonuçları çerçevesinde elde edilen doğruluk oranları karşılaştırıldığında etkin sonuçlar elde edildiği görülmüştür. Bulgular bölümünde KNN, SVM, MLP, RF ve derin öğrenme modelleri ile elde edilen test sonuçları hata matrisleri, metrikler ve roc eğrileri ile verilmiştir.

Aktivitelerin doğruluk oranları, modelden elde edilen genel doğruluk oranı (accuracy) ve kullanılan algoritmaların karşılaştırma sonuçları çizelgeler halinde bu bölümde verilmiştir.

Çizelge 5.1. Aktiviteler ve algoritmaların %10 test verisi ile karşılaştırılması

Aktiviteler	%10 Test Verisi						
	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma	Accuracy
KNN	0,99	0,87	1,00	0,76	0,87	1,00	0,92
SVM	1,00	0,96	1,00	0,76	0,97	1,00	0,95
MLP	1,00	0,98	1,00	0,76	0,98	1,00	0,95
RF	1,00	0,55	0,98	0,74	0,88	1,00	0,87
Deep Learning	1,00	1,00	0,99	0,77	0,92	1,00	0,95

Çizelge 5.1’de görüldüğü üzere %10 test verisi ile yapılan uygulamada en düşük doğruluk oranları oturma aktivitesinde elde edilmiştir. Oturma aktivitesinde en yüksek doğruluk oranı %77 ile derin öğrenme modeli ile elde edilmiştir. Tırmanma aktivitesinde RF modelinin %55 ile en düşük doğruluk oranını verdiği görülmektedir. Test uygulamasında yürüme, uzanma ve inme aktivitelerinde modeller %98 ile %100 arasında yüksek başarı

sağlamışlardır. Modellerin genel doğruluk oranlarına bakıldığında SVM, MLP ve derin öğrenme modellerinin %95 başarı sağladığı görülmektedir. En düşük genel başarı oranı %87 ile RF modelinde elde edilmiştir.

Çizelge 5.2. Aktiviteler ve algoritmaların %20 test verisi ile karşılaştırılması

	%20 Test Verisi						
Aktiviteler	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma	Accuracy
KNN	0,84	0,78	0,98	0,88	0,68	1,00	0,86
SVM	0,88	0,66	1,00	0,97	0,37	1,00	0,81
MLP	0,98	0,72	1,00	0,99	0,26	1,00	0,82
RF	0,77	0,40	0,99	0,99	0,19	1,00	0,72
Deep Learning	0,95	0,83	1,00	0,55	0,91	1,00	0,87

Çizelge 5.2’de aktiviteler ve %20 test verisi ile elde edilen sonuçların modellere göre karşılaştırılması yapılmıştır. Sonuçlara göre %10 test verisine göre daha düşük doğruluk oranları elde edildiği görülmektedir. En düşük doğruluk oranları %19 ile RF modelinde bekleme aktivitesinde ve %40 ile tırmanma aktivitesinde elde edilmiştir. Modeller uzanma ve inme aktivitesinde yüksek başarı sağlamıştır ve doğruluk oranları %98 ile %100 arasında değişmektedir. Derin öğrenme modeli oturma aktivitesinde %55 doğruluk oranı verirken yürüme, tırmanma, inme, bekleme ve uzanma aktivitelerinde başarı oranı %83 ile %100 arasında değişmektedir. Genel doğruluk oranlarına bakıldığında en iyi sonucu %87 ile derin öğrenme modelinin verdiği, en düşük sonucun ise %72 ile RF modelinde elde edildiği görülmektedir.

Çizelge 5.3. Aktiviteler ve algoritmaların %30 test verisi ile karşılaştırılması

	%30 Test Verisi						
Aktiviteler	Yürüme	Tırmanma	İnme	Oturma	Bekleme	Uzanma	Accuracy
KNN	0,86	0,82	0,95	0,94	0,71	1,00	0,88
SVM	0,87	0,83	0,99	0,98	0,51	1,00	0,85
MLP	0,86	0,82	1,00	1,00	0,23	1,00	0,80
RF	0,81	0,55	0,95	0,99	0,15	1,00	0,73
Deep Learning	1,00	0,86	0,96	0,78	0,91	1,00	0,92

Çizelge 5.3' te %30 test verisi ile elde edilen sonuçlar aktivitelerle birlikte verilmiş ve algoritmalar karşılaştırılmıştır. %20 test verisinden daha yüksek doğruluk oranları elde edildiği görülmektedir. En düşük doğruluk oranları RF modeli ile bekleme aktivitesinde %15 olarak, MLP modeli ile bekleme aktivitesinde %23, RF modelinde tırmanma aktivitesinde %55, SVM modelinde bekleme aktivitesinde %51 olarak görülmektedir. Uzanma aktivitesi tüm modellerde %100 doğruluk oranı ile tahmin edilmiştir. Yürüme aktivitesinde %81 ile %100, inme aktivitesinde %95 ile %100, oturma aktivitesinde %78 ile %100 arasında değişen doğruluk oranları elde edilmiştir. Genel doğruluk oranlarına bakıldığında %92 doğruluk oranı ile en başarılı algoritma derin öğrenme modeli olurken, en düşük doğruluk oranı %73 ile RF modelinde elde edilmiştir.

Çalışmada %10 test verisi ile en yüksek doğruluk oranlarına ulaşıldığı görülmektedir. Sensör verileriyle aktivite tanımlamada makine öğrenmesi modelleri veya derin öğrenme modelleriyle yapılacak çalışmalarda daha büyük veri setleriyle daha iyi sonuçların elde edilebileceği düşünülmektedir. Elde edilen sonuçlarda modellerin bazı aktiviteleri tahmin etmede başarı oranının düşük olduğu görülmüştür. Bunun nedeni kullanılan veri setindeki aktivitelere ait kayıtların sayısının yetersiz olması olabilir.

Bu kapsamda SVM, MLP ve deep learning modellerinin sunucu ortamına dağıtımının yapılması halinde, anlık olarak güvenlik personellerinin giyilebilir teknolojilerle veya taşıyacakları akıllı telefonlarla aktivitelerinin tanımlanabileceği sonucuna ulaşılmıştır.

6.KAYNAKLAR

- Alsheikh, M. A., Niyato, D., Lin, S., Tan, H. P., & Han, Z. (2016). Mobile Big Data Analytics Using Deep Learning and Apache Spark. *IEEE Network*, 30, 3.
- Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics Conference Series*.
- Aydemir, M. S. (2019). Hiperspektral Görüntülerde Yarı GÜdümlü Öğrenme Teknikleri. İstanbul: Yıldız Teknik Üniversitesi.
- Balkır, U. B. (2020, 1 22). Ulusal Tez Merkezi. İSTANBUL: YÖK Tez Merkezi. YÖK Tez Merkezi. adresinden alındı
- Bao, L., & Intille, S. S. (2004). Activity Recognition from User-Annotated Acceleration Data. *International Conference on Pervasive Computing* (s. 1-17). Berlin: Pervasive.
- Berchtold, M., Budde, M., Gordon, D., Schmidtke, H. R., & Beigl, M. (2010). ActiServ: Activity Recognition Service for Mobile Phones. *Proceedings - International Symposium on Wearable Computers* (s. 1-8). ISWC.
- Berrar, D. (2019). Cross-validation. *Reference Module in Life Sciences*, 542-545.
- Breiman, L. (2001, 1). *Random Forests*. stat.berkeley.edu: <https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf> adresinden alındı
- Datasciencelovers. (2020, 3 28). K-Nearest Neighbors (KNN) – Theory. datasciencelovers.com: <http://www.datasciencelovers.com/machine-learning/k-nearest-neighbors-knn-theory/> adresinden alındı
- Görmez, B. (2021). Adli Bilişimde Makine Öğrenmesi: Makine Öğrenmesi Algoritmaları İle Terör Olaylarının Tahmin Edilmesi Çalışması. Ankara: Ankara Üniversitesi.
- Györi, N., Fabian, A., & Homanyi, G. (2009). An Activity Recognition System For Mobile Phones. *Mobile Networks and Applications* (s. 82-91). Berlin: ACM.
- Hastie, T., Tibshirani, R., & Friedman, J. (2008, 8). *ESLII*. web.stanford.edu: <https://web.stanford.edu/~hastie/Papers/ESLII.pdf> adresinden alındı
- Hernandez, V., Suzuki, T., & Venture, G. (2020). Convolutional and recurrent neural network for human activity recognition: Application on American sign language. *PLoS One*, 15.
- Hoo, Z. H., Candlish, J., & Teare, D. (2017). What is an ROC curve? *Emergency Medicine*.
- Hyunh, P., & Yoo, M. (2016, 5 28). VLC-Based Positioning System for an Indoor Environment Using an Image Sensor and an Accelerometer Sensor. *Sensors*, s. 783.
- Jakkula, V. (2011, 9 5). SVM Tutorial. 11 8, 2021 tarihinde course.ccs.neu.edu: <https://course.ccs.neu.edu/cs5100f11/resources/jakkula.pdf> adresinden alındı
- Kern, N., Schiele, B., & Schmidt, A. (2003). Multi-sensor Activity Context Detection for Wearable Computing. *Ambient Intelligence* (s. 220-232). Veldhoven: EUSAI.
- Kesilmiş, Z. (2011). An artificial neural network controlled microelectromechanical accelerometer. Ulusal Tez Merkezi.

- Kuzu, A. (2006). Multisensor Fusion Of Mems Accelerometers For Measurement Accuracy Improvement. İstanbul: İstanbul Teknik Üniversitesi.
- Kwapisz, J., Weiss, G., & Moore, S. (2012, 12 2). *WISDM: Wireless Sensor Data Mining*. 1 3, 2022 tarihinde Wireless Sensor Data Mining: <https://www.cis.fordham.edu/wisdm/dataset.php> adresinden alındı
- Lara, O. D., & Labrador, M. A. (2013). A Survey on Human Activity Recognition using Wearable Sensors. *Communications Surveys & Tutorials* (s. 1192-1209). IEEE.
- Lee, S. W., & Mase, K. (2002). Activity and location recognition using wearable sensors. *Pervasive Computing* (s. 24-32). IEEE.
- Li, X., Marsic, I., Zhang, Y., & Yang, J. (2016). Deep Neural Network for RFID-Based Activity Recognition. *Wireless of the Students, by the Students, and for the Students*. NewYork: ACM MobiCom.
- Marin, G., Dominio, F., & Zanuttigh, P. (2014). Hand gesture recognition with Leap Motion and Kinect devices. *IEEE International Conference on Image Processing (ICIP)*. Paris: IEEE.
- Marius, P., Mastorakis, N. E., & Balas, V. E. (2009). Multilayer perceptron and neural networks. *Transactions on Circuits and Systems*. 8. WSEAS.
- Maurer, U., Smailagic, A., Siewiorek, D. P., & Deisher, M. (2006). Activity Recognition and Monitoring Using Multiple Sensors on Different Body Pozitions. *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks* (s. 116). Cambridge: IEEE.
- Ming, L. (2013, 3 10). A Study of Mobile Sensing Using Smartphones. *Ubiquitous Sensing: Services and Applications*.
- Mohammed, S. H. (2021). Detection of cancer area in lung images with the help of deep learning algorithms . Elazığ: Firat Üniversitesi.
- Oluwalade, B., Neela, S., Wawira, J., Adejumo, T., & Purkayastha, S. (2021). Human Activity Recognition using Deep Learning Models on Smartphones and Smartwatches Sensor Data. *14th International Joint Conference on Biomedical Engineering Systems and Technologies*. Viyana: HEALTHINF.
- Ortiz, J. R., Anguita, D., Ghio, A., Oneto, L., & Parra, X. (2021, 10 29). *UCI Machine Learning Repository*. [archive.ics.uci.edu: https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones](https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones) adresinden alındı
- Paksoy, T. (2021). Machine Learning Methods For Survival Data. İzmir: Dokuz Eylül Üniversitesi.
- Passaro, V., Cuccovillio, A., Vaiani, L., De Carlo, M., & Campanella, C. (2017). Gyroscope Technology and Applications: A Review in the Industrial Perspective. *Sensors*(2284).
- Patel, S. (2017, 5 3). Chapter 2 : SVM (Support Vector Machine) — Theory. medium.com: <https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72> adresinden alındı
- Rahimi, A., Hemingway, B., Landay, J. A., Choudhury, T., Consolvo, S., Harrison, B., . . . Haehnel, D. (2008). The Mobile Sensing Platform: An Embedded Activity Recognition System. *Pervasive Computing*. IEEE.

- Ranasinghe, S., Al Machot, F., & Mayr, H. C. (2016). A review on applications of activity recognition systems with regard to performance and evaluation. *International Journal of Distributed Sensor Networks*. 12. SAGE.
- Sakacı, B. (2018). Askeri Personel İçin Akıllı Kıyafet Tasarımı. İstanbul: Yıldız Teknik Üniversitesi.
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 160.
- Schott, M. (2019, 4 22). K-Nearest Neighbors (KNN) Algorithm for Machine Learning. medium.com: <https://medium.com/capital-one-tech/k-nearest-neighbors-knn-algorithm-for-machine-learning-e883219c8f26> adresinden alındı
- Schott, M. (2019, 4 25). Random Forest Algorithm for Machine Learning. medium.com: <https://medium.com/capital-one-tech/random-forest-algorithm-for-machine-learning-c4b2c8cc9> adresinden alındı
- Shrestha, A., & Mahmood, A. (2019). Review of Deep Learning Algorithms and Architectures. *IEEE Access*, 1-1.
- Singh, D., Merdivan, E., Psychoula, I., Kropf, J., Hanke, S., Geist, M., & Holzinger, A. (2017). Human Activity Recognition Using Recurrent Neural Networks. *Human Activity Recognition Using Recurrent Neural Networks* (s. 267-274). Laxenburg: Springer International.
- Şirin, E. (2017, 7 2). Hata Matrisini Yorumlama. veribilimiokulu.com: <https://www.veribilimiokulu.com/hata-matrisini-confusion-matrix-yorumlama/> adresinden alındı
- Taravat, A., Proud, S., Peronaci, S., Frate, F. D., & Oppelt, N. (2015). Multilayer Perceptron Neural Networks Model for Meteosat Second Generation SEVIRI Daytime Cloud Masking. *Remote Sens*, s. 1529-1539.
- veribilimcisi. (2021, 1 1). ROC Eğrileri. veribilimcisi.com: <https://veribilimcisi.com/2017/07/18/roc-egrisi-roc-curve/> adresinden alındı
- Weiss, G. M., & Lockhart, J. W. (2012). The Impact of Personalization on Smartphone-Based Activity Recognition . *Innovative Applications of Artificial Intelligence* (s. 1541-1546). North America: AAAI.
- Yang, J. Y., Chen, Y. P., Lee, G. Y., Liou, S. N., & Wang, J. S. (2007). Activity Recognition Using One Triaxial Accelerometer: A Neuro-fuzzy Classifier with Feature Reduction. *Entertainment Computing* (s. 395-400). Berlin: ICEC.
- Yiu, T. (2019, 6 12). Understanding Random Forest. towardsdatascience.com: <https://towardsdatascience.com/understanding-random-forest-58381e0602d2> adresinden alındı

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : : T.C.
 Uyuşu :
 Doğum tarihi ve yeri :
 Medeni hali :
 Telefon :
 Faks :
 e-mail :

Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Yüksek lisans	SBTÜ/ Savunma Teknolojileri	Devam ediyor
Yüksek lisans	GOP Üniversitesi/ Eğitim Bilimleri	2018
Lisans	Cumhuriyet Üniversitesi/Bilgisayar Mühendisliği	2019
Lisans	Kocaeli Üniversitesi/Bilgisayar Öğretmenliği	2003
Lise	Tokat Mesleki ve Teknik Anadolu Lisesi	1998

İş Deneyimi

Yıl	Yer	Görev
2021-Halen	Burdur İl Milli Eğitim Müdürlüğü	Bilgisayar Mühendisi
2009-2021	Tokat Yeşilırmak MTAL	Bilgisayar Öğretmeni
2007-2009	Tokat Fatih İlköğretim Okulu	Bilgisayar Öğretmeni
2003-2007	Yeşilyurt Milli Hakimiyet İ.O/TOKAT	Bilgisayar Öğretmeni

Yabancı Dil

İngilizce

Yayınlar

Hobiler

Yüzme, Fitness, Kamp,Şii



**SİVAS
BİLİM VE TEKNOLOJİ
ÜNİVERSİTESİ**

KÖKLERDEN GÖKLERE...