



# Jiroskop ve İvme Ölçer Verileriyle Makine Öğrenmesi Algoritmaları Kullanılarak İnsan Aktivitesi Tanımlama

Cem Özkaya<sup>1\*</sup>, Mustafa Yasin Esas<sup>2</sup>

<sup>1\*</sup> TOBB Eko. ve Tek. Üniversitesi, Fen Bilimleri Enstitüsü, Biyomedikal Mühendisliği Bölümü, Ankara, Türkiye, (ORCID: 0000-0002-5029-6941), [cozkaya@etu.edu.tr](mailto:cozkaya@etu.edu.tr)

<sup>2</sup> Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Biyomedikal Mühendisliği Bölümü, Kayseri, Türkiye, (ORCID: 0000-0002-2677-5897), [mustafaesas@gmail.com](mailto:mustafaesas@gmail.com)

(1st International Conference on Applied Engineering and Natural Sciences ICAENS 2021, November 1-3, 2021)

(DOI: 10.31590/ejosat.1011461)

**ATIF/REFERENCE:** Özkaya, C. & Esas, M. Y. (2021). Jiroskop ve İvme Ölçer Verileriyle Makine Öğrenmesi Algoritmaları Kullanılarak İnsan Aktivitesi Tanımlama. *European Journal of Science and Technology*, (28), 811-818.

## Öz

Günümüzde akıllı cihazlar aracılığıyla (telefon, saat vb.) fiziksel hareket ve yakılan kalori miktarının takip edilmesi cihaz kullanıcıları tarafından yaygın olarak kullanılmaktadır. Bu akıllı cihazlar vasıtasıyla elde edilen verilere dayalı olarak gerçekleştirilen aktivitenin yüksek doğrulukla tanımlanması bazı hastalıkların veya fizyolojik eksikliklerin ortaya çıkarılması ve tedavisinde fikir verebilmektedir. Bu kapsamda çalışmamızda akıllı telefon vasıtasıyla elde edilmiş olan jiroskop ve ivmeölçer verilerine dayalı olarak yüksek doğrulukla aktivite tanımlanması çalışması gerçekleştirilmiştir. Yapılan çalışmada aktivite tanımlaması olarak merdiven inme, merdiven çıkma, koşma ve ayakta durma verileri sınıflandırılmıştır. Bu verileri ait on iki adet öznelik tespit edilmiş ve bu öznelikler sınıflandırma algoritmalarında kullanılmıştır. Çalışma, MatLab "Classification Learner Toolbox" uygulamasında bulunan sınıflandırma yöntemleri kullanılarak gerçekleştirilmiştir. Elde edilen bulgular literatür ile uyumlu ve katkı sağlayacak niteliktedir. Çalışmada hareket tanımlaması kapsamında %90 seviyelerinde doğruluk oranları elde edilmiştir. Çalışmamız ön araştırma niteliğinde olup, sonraki çalışmalarımız için faydalı bulgular içermektedir.

**Anahtar Kelimeler:** İnsan aktivitesi tanımlama, Jiroskop, İvmeölçer, Aktivite sınıflandırma, Sınıflandırma yöntemleri, Makine öğrenmesi

## Identifying Human Activity Using Machine Learning Algorithms with Gyroscope and Accelerometer Data

### Abstract

Today, tracking physical movement and the amount of calories burned through smart devices (phone, watch, etc.) is widely used by device users. Identifying the activity performed with high accuracy based on the data obtained by these smart devices can give an idea in the detection and treatment of some diseases or physiological deficiencies. In this context, in our study, a high-accuracy activity identification study was carried out based on the gyroscope and accelerometer data obtained via a smart phone. In the study, the data of stair down, stair up, running and standing were classified as activity definitions. Twelve features of this data were identified and these features were used in classification algorithms. The study was carried out using the classification methods available in the MatLab "Classification Learner Toolbox" application. The obtained findings are compatible with the literature and contribute to it. In the study, accuracy rates of 90% were obtained within the scope of motion identification. Our study is a preliminary research and contains useful findings for our further studies.

**Keywords:** Human activity identification, Gyroscope, Accelerometer, Activity classification, Classification methods, Machine learning

\* Corresponding Author: [cozkaya@etu.edu.tr](mailto:cozkaya@etu.edu.tr)

## 1. Giriş

Çalışmamızda akıllı telefon içerisinde bulunan sensörler kullanılarak elde edilmiş üç eksenli ivmeölçer ve jiroskop verileri kullanılmıştır [1]. Veriler M. Malekzadeh ve arkadaşlarının yaptığı çalışmalar ile elde edilerek "kaggle.com/malekzadeh/motionsense-dataset" veri tabanında kullanıma sunulmuştur. Çalışmada amacımız veri setindeki jiroskop ve ivmeölçer verileri kullanılarak hareket tipinin yüksek doğrulukla tanımlanmasının sağlanmasıdır. Bu amaçla sınıflandırma çalışması yapılmıştır. Sınıflandırma çalışmasında makine öğrenmesi algoritmaları kullanılmış ve yüksek doğruluk oranları elde edilmiştir. Sonuçlar bölümünde kullanılan algoritmaların performansları değerlendirilip karşılaştırılmıştır. Literatürde insan aktivitesi tanımlanması üzerine birçok çalışma bulunmaktadır.

Wang ve arkadaşları [2] sağlık hizmetlerinde insan aktivitesi tanımlama amacıyla giyilebilir sensör kullanımı üzerine bir araştırma çalışmasında bulunmuşlardır. Sensör tabanlı insan etkinliği tanıma (HAR), yaşlı insanların günlük yaşamını desteklemek, fitness sistemleri, fabrika çalışanlarının izlenmesi vb. için en umut verici yardımcı teknolojilerden biridir ve insan merkezli uygulamalarda muazzam bir potansiyel sağlamıştır. Araştırmada yapılan çalışmada, yeni başlayanlar ve araştırmacılar için HAR'a daha kapsamlı bir bakış açısı sağlamak amaçlanmıştır. Hem geleneksel yaklaşımlar hem de derin öğrenme yöntemleri dahil olmak üzere sensör çeşitleri, sensör yerleşimi, günlük aktivitelerin tanımlanması, veri ön işleme yöntemleri (filtreleme, pencereleme vb.), öznelik belirleme, öznelik azaltma (PCA, Kernel PCA, Auto encoder, Sparse filtering, LDA vb.) ve öznelik seçme (filter methods, wrapper methods, embedded methods), sınıflandırma algoritmaları, performans değerlendirme yöntemleri, insan aktivitesi tanımlanmasının sağlık üzerine uygulamaları üzerinde detaylı olarak inceleme yapılmıştır. Bu konuda çalışma yapacak kişilere kılavuz oluşturulmuştur.

Jia ve Chen [3] sensör verileri kullanarak insan aktivitesi tanımlama çalışmasında bulunmuşlar ancak geleneksel yöntemlerin haricinde sadece sensör verisi değil bilgi ile entegre sensör verisi içeren bir tanımlama çalışmasında bulunmuşlardır. Daha spesifik olarak bu yaklaşım hiyerarşik olarak insan aktivitesiyle birlikte düşük seviyeli aksiyonları ve mimiklerin bileşimini ele almaktadır. Daha sonra bu hiyerarşik yapı, makine öğrenmesi yöntemlerini kullanarak otomatik akıl yürütmeli mantıksal formlere ve kural tabanlı yapılara dönüştürülmüştür. Bu çalışma video tabanlı insan etkinliği tanıma hakkında bazı açık kaynaklı veriler kullanılarak doğrulanmıştır. İnsan aktivitesi tanımlamada makine öğrenmesi ve sembolik akıl yürütmenin entegrasyonunun umut verici bir yöntem olduğundan bahsedilmiştir.

Nweke ve arkadaşları [4] mobil ve giyilebilir sensörler kullanılarak insan aktivitesi tanıma için derin öğrenme algoritmaları konusunda yapılan çalışmaları inceleyen bir araştırmada bulunmuşlardır. Çalışmada insan etkinliği tanıma sistemleri; ortam destekli yaşam, spor yaralanmalarının tespiti, yaşlı bakımı, rehabilitasyon ve akıllı ev ortamlarında eğlence ve gözetim alanındaki insan davranışlarının sürekli olarak izlenmesini sağlayan bir çerçevenin parçası olarak tanımlanmıştır. İlgili özneliklerin çıkarılması, mobil ve giyilebilir sensör tabanlı insan etkinliği tanıma hattının en zorlu kısmı olduğundan bahsedilmiştir. Öznelik çıkarmanın, algoritma performansını etkilediğinden ve hesaplama süresini ve karmaşıklığı

azalttığından bahsedilmiştir. Derin öğrenmenin ortaya çıkması ve artan hesaplama güçleriyle sınıflandırmada çeşitli alanlarda otomatik öznelik öğrenme için ve son zamanlarda mobil cihazlarda basit ve karmaşık insan etkinliği tanımının öznelik çıkarma ve sınıflandırması için derin öğrenme yöntemlerinin benimsendiğinden bahsedilmiştir. Öznelik öğrenimi için mobil veya giyilebilir sensörlerin ve derin öğrenme yöntemlerinin birleştirilmesinin çeşitlilik sağladığından, daha yüksek genelleme sunduğundan ve insan etkinliğinin tanınmasındaki zorlu sorunları ele aldığından bahsedilmiştir. Bu incelemenin odak noktası, mobil ve giyilebilir sensör tabanlı insan etkinliği tanıma için derin öğrenme yöntemlerinin derinlemesine özetlerini sağlamaktır. İnceleme yöntemleri, benzersizliği, avantajları ve sınırlamaları göstermektedir. Bu araştırma; çalışmaları yalnızca üretken, ayırt edici ve hibrit yöntemler olarak sınıflandırmakla kalmamakta, aynı zamanda birbirlerine göre avantajlarını da vurgulamaktadır. Çalışmada ayrıca daha fazla araştırma ve iyileştirme gerektiren araştırma problemleri de özetlenmiştir.

Subasi ve arkadaşları [5] akıllı telefon sensörlerinden elde edilen verilerin Bagging ve Boosting tekniklerini kullanılarak sınıflandırılma yapılmak üzere insan aktivite tanımlaması çalışmasında bulunmuşlardır. Bu çalışmada da sağlık uygulamalarında mobil teknolojilerin kullanımı, özellikle yaşlılar ve tıbbi kontrol altındakiler için günlük klinik aktivitelerin akıllı ve otomatik izlenmesi, uzaktan yaşam asistanları ve önleyici bakım için hekimleri ve hastaları bir araya getirdiği vurgulanmış ve insan aktivitesi tanımlamasının ileri ki zamanlarda yaşamın ayrılmaz bir parçası haline geleceğinden bahsedilmiştir. Makalenin amacı, Bagging ve Adaboost sınıflandırıcılarını kullanarak akıllı telefon sensörlerinin verilerine dayalı bir insan aktivitesi tanımlama sistemi geliştirmektir. İnsan aktivitesi verileri için deneysel sonuçlar, farklı veri madenciliği teknikleri uygulandıktan sonra değerlendirilmiştir. Her konu için accuracy, F1-Score ve ROC alanı hesaplanmıştır. Adaboost sınıflandırıcıları algoritmasının, akıllı telefon tabanlı insan aktivitesi tanımlamanın performansını önemli ölçüde iyileştirdiği gözlemlenmiş ve SVM ile birlikte diğer sınıflandırıcılara kıyasla %97,44 doğruluğa ulaştığı belirtilmiştir. Adaboost SVM'nin önerilen algoritması, sürekli bakıma ihtiyaç duyan yaşlı ve engelli hastalar için doğru bir insan aktivitesi tanımlamayı sağlamakla birlikte, tüm sağlık profesyonellerinin kararlarını destekleyen bir araç olduğundan bahsedilmiştir.

Literatürde insan aktivitesi tanımlanması üzerine çeşitli çalışmalar bulunmakta olup %90 üzerinde doğruluk oranları ile aktivite tanımlaması gerçekleştirilen sonuçlar gözlemlenebilmektedir.

## 2. Materyal ve Metot

Çalışmada kullanılan veriler daha önce kontrollü bir şekilde kaydedilmiş ve veriler üzerinde çeşitli akademik çalışmalar gerçekleştirilmiştir [1]. Jiroskop ve ivme ölçer verileri üç eksenli olup (x,y,z) toplam altı kanal sensör verisi bulunmaktadır. Bu çalışma Matlab R2021a uygulaması kullanılarak yapılmıştır.

Toplam 24 kişiye ait veri kullanılmış olup bu veriler 4 farklı fiziksel hareket gerçekleştirilirken 30 'ar saniye kayıt alınarak elde edilmiştir. Kullanılan fiziksel hareketler merdiven inme, merdiven çıkma, ayakta durma ve koşma verileridir. 24 kişinin her biri için hem jiroskop hem de ivmeölçer üzerinden 30 saniye merdiven inme, 30 saniye merdiven çıkma, 30 saniye ayakta durma ve 30 saniye koşma verisi 50 Hz örnekleme frekansı ile elde edilmiştir.

Veriler elde edilip derlendikten sonra bu verilere ait on iki adet öznelik oluşturulmuştur. Bu öznelikler temel istatistik parametreleri olan; ortalama (mean) (1. öznelik), karekök ortalama (RMS) (2. öznelik), standart sapma (standart deviation) (3. öznelik), basıklık (kurtosis) (4. öznelik), çarpıklık (skewness) (5. öznelik), maksimum ve minimum arasındaki fark (6. öznelik), değişim (variance) (7. öznelik), sinyal gürültü oranı (SNR) (8. öznelik), sinyal gürültü ve bozulma oranı (9. öznelik), karelerin kök toplamı seviyesi (root sum of squares level) (10. öznelik), ortalama frekans (11. öznelik) ve tepeden karekök ortalaması (peak to RMS) (12. öznelik) dir. Bu özneliklere ait formülasyonlar ve açıklamaları aşağıda sunulmuştur.

## 2.1. Öznelikler

### 2.1.1. Ortalama (mean)

$$\bar{x} = \frac{1}{N} \left( \sum_{i=1}^N x_i \right)$$

Aritmetik ortalama, elimizdeki verilerin değerlerinin toplamlarının o verilerin sayısına bölünerek elde edilen değerdir.

### 2.1.2. Karekök Ortalama (RMS)

$$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$$

Bir veri topluluğunun RMS değeri, bu topluluktaki her bir değer in karesinin, toplam veri sayısına olan oranının kareköküdür.

### 2.1.3. Standart Sapma

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Standart sapma, olasılık ve istatistiki olarak, veri setinde bulunan değerlerin dağılımını gösteren bir ölçüttür.

### 2.1.4. Basıklık (Kurtosis)

$$K = \frac{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^4}{\left( \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2} \right)^2}$$

Basıklık, bir dağılımın sivrilik derecesidir ve genellikle normal dağılıma göre ele alınır. Normal dağılım, simetrik dağılımın özel bir halidir. Bir dağılımın normal olabilmesi için hem simetrik olması hem de normal bir yüksekliğe sahip olması gerekir.

### 2.1.5. Çarpıklık (Skewness)

$$S = \frac{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^3}{\left( \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2} \right)^3}$$

Çarpıklık, örnek ortalama etrafındaki verilerin asimetrisinin bir ölçüsüdür. Çarpıklık negatifse, veriler sağdan ziyade ortalamanın soluna yayılır. Çarpıklık pozitifse, veriler daha çok sağa doğru yayılır.

### 2.1.6. Max-Min Arasındaki Fark

$$x = \{x_1, x_2, x_3, \dots, x_N\}$$

$$, x_{(max-min)} = x_{max} - x_{min}$$

Tepe değeri, herhangi bir yöndeki maksimum genlik bilgisinin mutlak değeridir. Diğer bir deyişle pozitif yöndeki maksimum değerler ile negatif yöndeki maksimum değer arasındaki fark, tepe değeri olarak adlandırılır.

### 2.1.7. Değişim (Variance)

$$\sigma^2 = \left( \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)$$

Olasılık ve istatistiki olarak varyans, mümkün bütün değerlerin beklenen değer veya ortalamadan uzaklıklarının karelerinin ortalaması şeklinde bulunan bir ölçüdür.

### 2.1.8. Sinyal Gürültü Oranı (SNR)

$$SNR = \frac{\bar{X}^2}{\sigma^2}$$

SNR, bir verinin ortalamasının karesinin standart sapmasının karesine (değişim) oranıdır.

### 2.1.9. Sinyal Gürültü ve Bozulma Oranı (SINAD)

$$THD = \sqrt{\frac{H_1^2 + H_2^2 + H_3^2 + H_4^2 + H_5^2}{F}}$$

$$SINAD = -10 \log \left[ 10^{-\frac{SNR}{10}} + 10^{-\frac{THD}{10}} \right]$$

Sinyal genliğinin diğer spektral bileşenlerin toplamına oranıdır. Spektral bileşenler beş adet harmonik frekansları içerir. SINAD toplam harmonik bozulma (THD) ve gürültüye eş değerdir.

### 2.1.10. Karelerin Kök Toplamı Seviyesi (Root Sum of Squares Level)

$$x = \{x_1, x_2, x_3, \dots, x_N\}$$

$$RSS = \sqrt{X_1^2 + X_2^2 + X_3^2 \dots \dots \dots + X_N^2}$$

RSS, istatistiksel bir tolerans analiz yöntemidir. Çoğu durumda, gerçek bireysel parça boyutları, tolerans sınırlarına yakın gerçek boyutlara sahip çok az parça ile tolerans aralığının merkezine yakın bir yerde meydana gelir.

### 2.1.11. Ortalama Frekans

$$f_{ort} = \frac{\sum_{i=0}^N I_i f_i}{\sum_{i=0}^N I_i}$$

N = spektrumdaki frekans kutusu sayısı

$f_i$  = N'nin i kutusundaki spektrum frekansı

$I_i$  = N'nin i kutusunda spektrumun yoğunluğu

Bir spektrumun ortalama frekansı, spektrogram yoğunluğunun (dB cinsinden) ve frekansın çarpımının toplamının, spektrogram yoğunluğunun toplamına bölünmesiyle hesaplanır.

### 2.1.12. Tepeden Karekök Ortalaması (Peak to RMS)

$$X_{prms} = \frac{x_{(max-min)}}{2\sqrt{2}}$$

Tepe değeri herhangi bir yöndeki maksimum genlik varyasyonunun mutlak değeridir. Pozitif yöndeki maksimum değerler ile negatif yöndeki maksimum değer arasındaki fark, tepe değeri olarak adlandırılır. RMS genlik, alternatif bir sinyalin ortalama genliğini temsil etmek için türetilmiş bir genliktir.

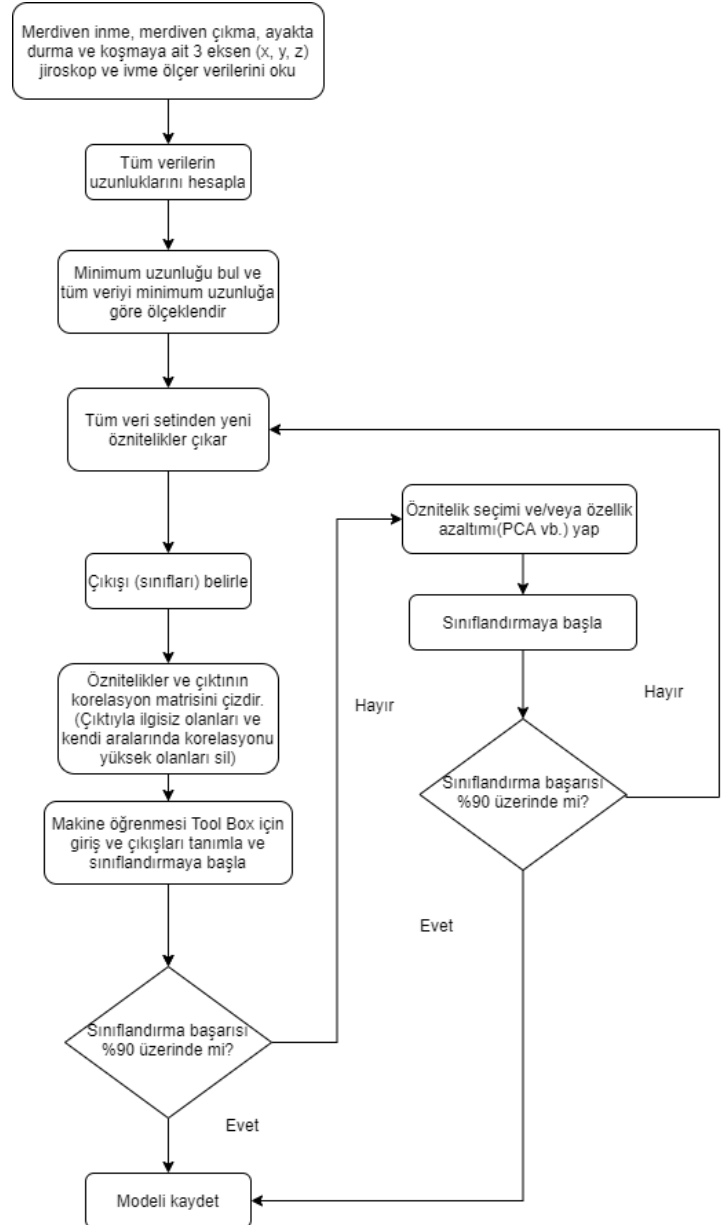
Burada elde edilen öznelik değerlerinin kendi aralarında ve çıktıyla olmak üzere korelasyon tablosu oluşturulmuştur. Korelasyon tablosunu oluşturmada iki temel husus bulunmaktadır. Burada öncelikle her bir öznelik ile çıktı arasındaki korelasyonlar incelenmiştir. Eğer çıktı ve öznelik arasındaki korelasyon sıfır veya sıfıra yakınsa tespit edilen öznelik modelimizin sınıflandırma başarısını düşüreceğinden ve hesaplama süresini uzatacağından modelimizden çıkarılmıştır. Diğer bir husus ise özneliklerin kendi arasındaki korelasyonları incelemektir. Eğer iki öznelik arasında yüksek bir korelasyon var ise burada çıktıyla korelasyonu daha düşük olan öznelik modelimizden çıkarılmıştır.

Elde edilen veriler, bu verilerden yola çıkarak oluşturulan öznelikler ve bu özneliklerin uygun olanları seçildikten sonra Matlab R2021a "Classification Learner" uygulaması kullanılarak bu uygulamada bulunan yirmi dokuz adet sınıflandırma algoritması denenmiştir. Bu uygulamada karar ağaçları, discriminant analizleri, logistic regression, naive bayes, destek vektör makineleri, k en yakın komşu, ensemble ağlar, sinir ağları ve bu yöntemlerin alt dalları bulunmaktadır. Veri setinin %80'ni eğitim için %20'si test için kullanılmıştır. Bu uygulamada bulunan tüm sınıflandırma algoritmaları beş defa çalıştırılmıştır. Algoritmalarından elde edilen başarı ortalamaları alınmıştır.

Yapılan araştırmalarda makine öğrenmesi teknikleri ile gerçekleştirilen insan aktivite tanımlanması çalışmalarında karar ağaçları kullanıldığında %51 ile %80 arasında doğruluk oranları, k-NN kullanıldığında %90 düzeylerinde doğruluk oranları, SVM kullanıldığında ise %66 ile %100 arasında doğruluk oranları tespit edildiği görülmektedir [6].

Makine öğrenmesi modellerine ait olan en yüksek doğrulukta çıktılar, sonuçlar bölümünde sunulmuştur. Elde edilen çıktılarının doğruluk yüzdeleri karşılaştırılmıştır. Ayrıca sınıflandırma sonucunda hangi hareketin tanımlanmasının daha yüksek oranda doğrulukla tespit edildiğinin incelenmesi amacıyla "Karışıklık Matrisi" (Confusion Matrix) tespit edilmiştir. Ayrıca bu sonuçlara ait ROC eğrisi de sunulmuştur.

Yapılan çalışmaya ait ve bu bölümde anlatılanları bir anlamda özetleyen akış şeması "Şekil-1"de verilmiştir.



Şekil 1. Modele ait akış şeması

## 3. Araştırma Sonuçları ve Tartışma

İkinci bölümde açıklanan özneliklerin belirlenmesi ve bu özneliklerin matematiksel olarak çözülmesi işlemleri Matlab kod ortamında gerçekleştirilmiştir. Bu işlemler gerçekleştirildikten sonra 13 sütun (12 adet öznelik ve çıktı) ve 576 satır (24 kişi, 3 eksen (x,y,z), merdiven inme, merdiven çıkma, ayakta durma ve koşma olmak üzere 4 adet hareket sınıfı, jiroskop ve ivme ölçer 2 sensör) veri bilgisi elde edilmiştir. Elde edilen 12 öznelik ve çıktıya ait korelasyon tablosu "Şekil-2"de gösterilmiştir. Buradaki şekilde de görüleceği üzere 12. öznelik olan tepeden karekök ortalaması (peak to RMS) ile çıktı arasında bir korelasyon olmadığından bu öznelik modelimizden çıkarılmıştır. 2. öznelik ile 10. öznelik arasındaki korelasyon 1.00 değerinde olduğundan ayrıca 2. öznelik ile 3. öznelik arasındaki korelasyon 0.95 değerinde olduğundan 2. ve 10. öznelikler modelimizden çıkarılmıştır. Burada modelimizden çıkarılacak öznelik tercih edilirken çıktıyla olan korelasyonu düşük özneliklerden başlayarak öznelik çıkarımı yapılmıştır. 3.

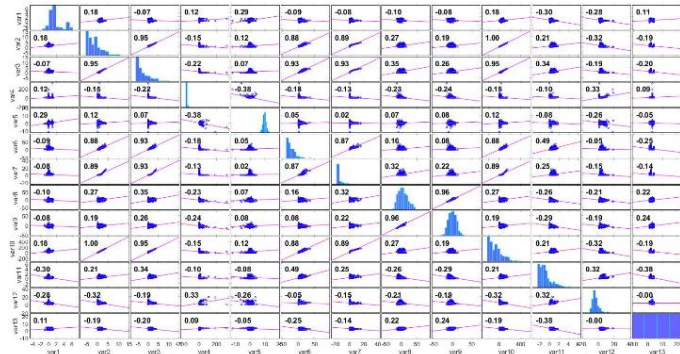
öz nitelik ile 6. öz nitelik arasındaki korelasyon 0.93 değerinde olduğundan ve bunların çıktıyla olan korelasyonu düşük olan öz nitelik 3. öz nitelik olduğundan bu öz nitelik de modelimizden çıkarılmıştır. Aynı şekilde 6. öz nitelik ile 7. öz nitelik arasındaki korelasyon 0.87 değerinde olduğundan ve bunların çıktıyla olan korelasyonu düşük olan öz nitelik 7. öz nitelik olduğundan bu öz nitelik de modelimizden çıkarılmıştır. Aynı şekilde 8. öz nitelik ile 9. öz nitelik arasındaki korelasyon 0.96 değerinde olduğundan ve bunların çıktıyla olan korelasyonu düşük olan öz nitelik 8. öz nitelik olduğundan bu öz nitelik de modelimizden çıkarılmıştır.

Makine Öğrenmesi Modeli	Sınıflandırma Sonucu (Doğruluk)
Destek Vektör Makineleri (DVM) (Medium Gaussian)	87.7 %
k-NN (Weighted k-NN)	87.3 %
Yapay Sinir Ağları (YSA) (Medium NN)	87.2 %

Sınıflandırma başarısını arttırmak için principle component analysis (PCA) gibi boyutsal öz nitelik çıkartım işlemleri de (dimensional feature reduction) yapılmaktadır. Çalışmamızda Matlab Classification Learner uygulamasında bulunan PCA kullanılarak sınıflandırma başarısını arttırmak hedeflenmiş ancak "Tablo-1"de elde edilen başarıdan üstün bir başarı elde edilmediğinden burada sunulmamıştır.

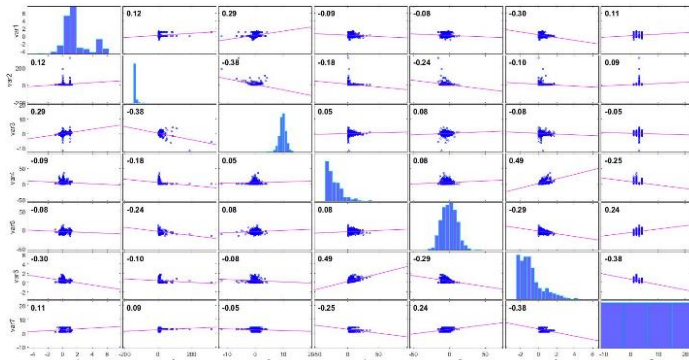
Aşağıda "Tablo-1"de elde edilen sınıflandırma başarılarına ait Karışıklık Matrisleri görülmektedir. 1-Merdiven İnme, 2-Koşma, 3-Ayakta Durma, 4-Merdiven Çıkma veri sınıflarını temsil etmektedir.

"Şekil-4"te görüldüğü üzere 2 ve 3 numaralı hareketler, 2 ve 4 numaralı hareketler, 3 ve 1 numaralı hareketler, 3 ve 2 numaralı hareketler yüksek doğruluk oranları ile DVM algoritması kullanılarak tespit edilmiştir.



Şekil 2. "12" adet öz nitelik ve çıktılarına ait korelasyon grafiği

2., 3., 7., 8., 10. ve 12. öz nitelikler modelimizden çıkarıldıktan sonra elde edilen korelasyon tablosu "Şekil-3"te gösterilmiştir. Ayrıca "Şekil-2"deki grafiğe bakıldığında öz niteliklerin dağılımları incelendiğinde herhangi bir aykırı (outlier) veri bulunmadığı değerlendirilmiştir. Bu sebeple veri setinden veri çıkartımı yapılmamıştır. Yine "Şekil-2"de görüleceği üzere veri setinden elde edilen dört sınıf çıktı (merdiven inme/çıkma, koşma ve ayakta durma) eşit olarak bulunduğu için veri dengelemesi yapılmamıştır.



Şekil 3. Diğer öz nitelikler çıkarıldıktan sonra elde edilen korelasyon grafiği (Modelimizde kullanılmak üzere seçilen öz nitelikler sırasıyla 1., 4., 5., 6., 9. ve 11.)

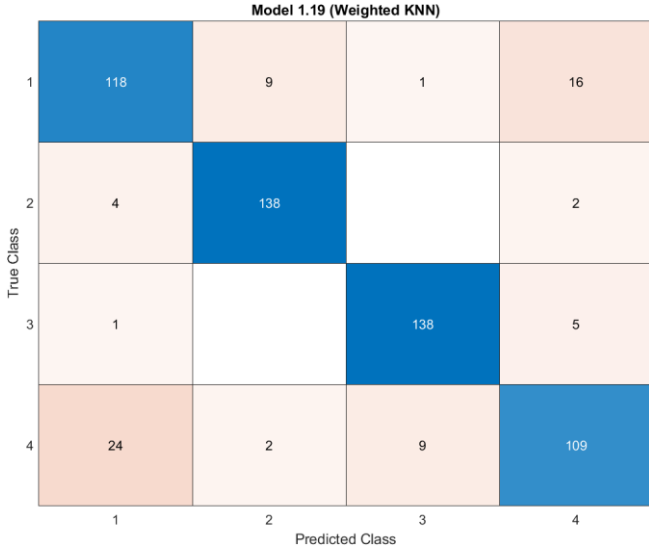
Öz niteliklerin seçimi yapıldıktan sonra Matlab R2021a "Classification Learner" uygulaması kullanılarak elde edilen en yüksek üç makine öğrenmesi algoritması "Tablo-1"de sunulmuştur. Burada görülen sonuçlar veri setinin %80'inin eğitim seti, %20'sinin test seti olması durumunda üç makine öğrenmesi algoritması ile elde edilen sınıflandırma başarıları görülmektedir.

		Model 1.12 (Medium Gaussian SVM)			
		1	2	3	4
True Class	1	113	6	1	24
	2	5	139		
	3			141	3
	4	26	3	3	112
		Predicted Class			

Şekil 4: DVM Karışıklık Matrisi (Tüm veri için yapılan çalışma)

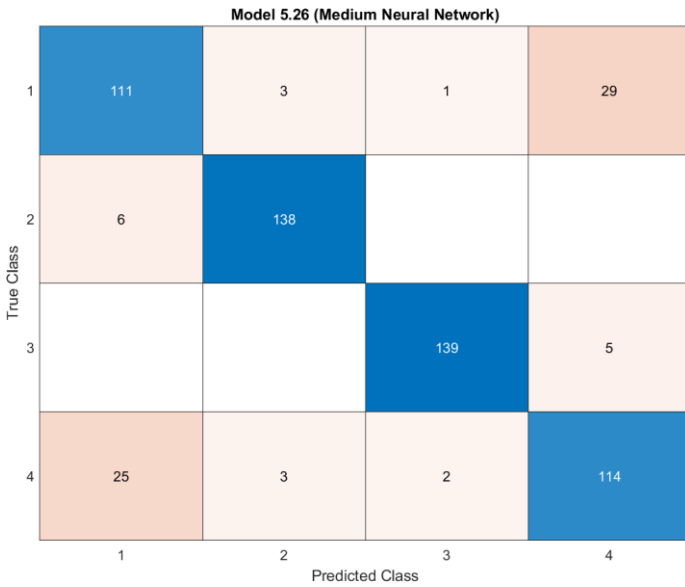
"Şekil-5"te görüldüğü üzere 2 ve 3 numaralı hareketler k-NN algoritmalarında yüksek doğrulukla tespit edilmiştir.

Tablo 1. Sınıflandırma Sonucu (Tüm veri için yapılan çalışma)



Şekil 5: k-NN Karışıklık Matrisi (Tüm veri için yapılan çalışma)

“Şekil-6”da görüldüğü üzere 2 ve 3 numaralı hareketler, 2 ve 4 numaralı hareketler, 3 ve 1 numaralı hareketler, 3 ve 2 numaralı hareketler YSA algoritmalarında yüksek doğrulukla tespit edilmiştir.



Şekil 6: YSA Karışıklık Matrisi (Tüm veri için yapılan çalışma)

Veri setinden elde edilen öznelik bilgisinden “x” eksen çıkarılarak aynı çalışmalar tekrar edilmiştir. Aşağıda görülen “Tablo-2”de “x” eksen çıkarılarak yapılan sınıflandırma çalışmasında edilen sınıflandırma başarı yüzdeleri görülmektedir. Burada görülmektedir ki bu eksen çıkarıldığında daha yüksek bir başarı elde edilememiştir. Burada “x” eksen bilgisinin sınıflandırma başarı oranını arttırmada katkısı olduğu söylenebilir.

Tablo 2. Sınıflandırma Sonucu (x eksen çıkarılmış olarak)

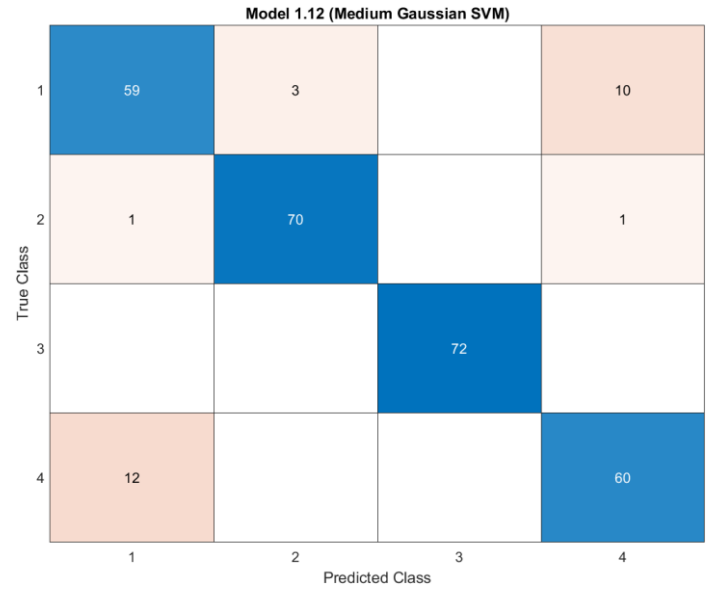
Makine Öğrenmesi Modeli	Sınıflandırma Sonucu (Doğruluk)
Destek Vektör Makineleri (DVM) (Qubic)	85.9 %
Yapay Sinir Ağları (YSA) (Wide NN)	85.7 %
k-NN (Fine)	84.9 %

Veri setinden elde edilen öznelik bilgisinden sadece jiroskoptan elde edilen veriler kullanılarak aynı çalışmalar tekrar edilmiştir. Aşağıda görülen “Tablo-3”te yapılan sınıflandırma çalışmasında edilen sınıflandırma başarı yüzdeleri görülmektedir. Burada görülmektedir ki elimizdeki veri seti ve çıkartılan öznelikler değerlendirildiğinde sadece jiroskop verilerini kullanarak sınıflandırma başarı oranı artmaktadır.

Tablo 3. Sınıflandırma Sonucu (sadece jiroskop verileri kullanılarak)

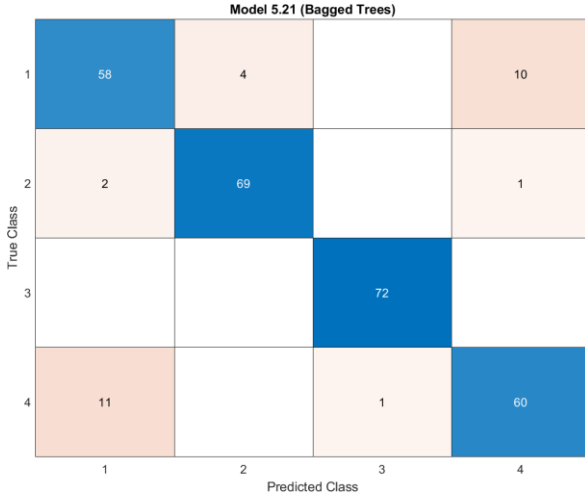
Makine Öğrenmesi Modeli	Sınıflandırma Sonucu (Doğruluk)
Destek Vektör Makineleri (DVM) (Medium Gaussian)	90.6 %
Ensemble (Bagged Trees)	89,9 %
k-NN (Weighted k-NN)	89.6 %

“Şekil-7”de görüldüğü üzere 1 ve 3 numaralı hareketler, 2 ve 3 numaralı hareketler, 3 ve 4 numaralı hareketler, 4 ve 2 numaralı hareketler yüksek doğruluk oranları ile DVM algoritması kullanılarak tespit edilmiştir.



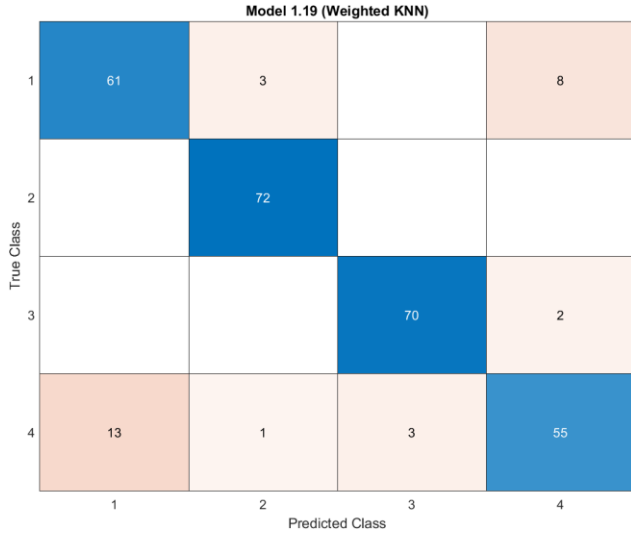
Şekil 7: DVM Karışıklık Matrisi (Jiroskop verisi için yapılan çalışma)

“Şekil-8”de görüldüğü üzere 1 ve 3 numaralı hareketler, 2 ve 3 numaralı hareketler, 3 ve 4 numaralı hareketler, 2 ve 2 numaralı hareketler yüksek doğruluk oranları ile Ensemble (Bagged Trees) algoritması kullanılarak tespit edilmiştir.

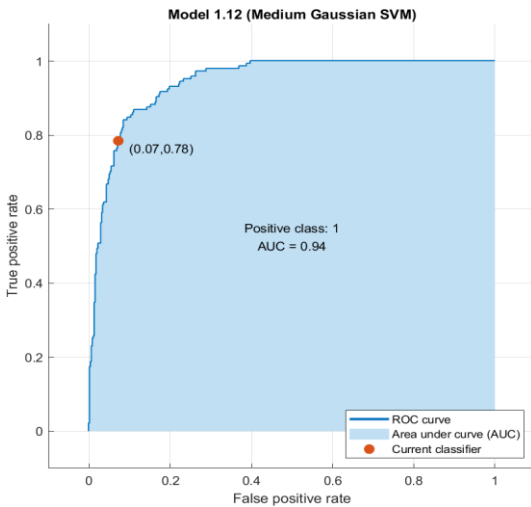


Şekil 8: Ensemble (Bagged Trees) Karışıklık Matrisi (Jiroskop verisi için yapılan çalışma)

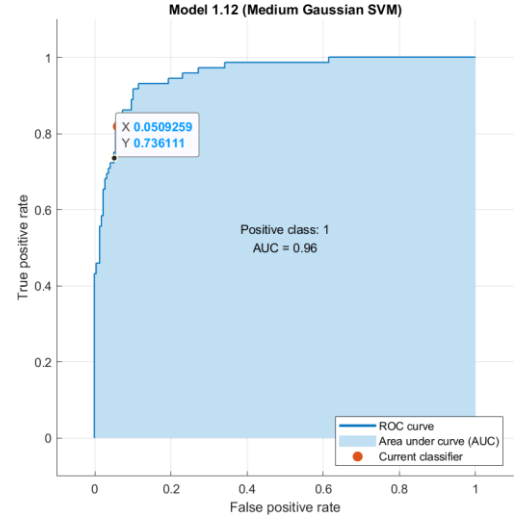
Şekil-9’da görüldüğü üzere 1 ve 3 numaralı hareketler, 2 ve 3 numaralı hareketler, 2 ve 4 numaralı hareketler yüksek doğruluk oranları ile k-NN algoritması kullanılarak tespit edilmiştir.



Şekil 9: k-NN Karışıklık Matrisi (Jiroskop verisi için yapılan çalışma)



Şekil 10: ROC-AUC Eğrisi (Tüm veri için yapılan çalışma)



Şekil 11: ROC-AUC Eğrisi (Jiroskop verisi için yapılan çalışma)

Bu çalışmada seçilecek modelde en az %90 başarı oranı yakalamak amacıyla çeşitli uygulamalar yapılmıştır. Uygulamalardan tercih edilen birinci yöntem jiroskop ve ivmeölçer verisinden elde edilen öznitelikler kullanılarak modeller denenmiştir. Bu işlem sonucunda “Tablo-1” ve “Şekil-4”te görüleceği üzere %87.7 doğrulukla sınıflandırma yapılmaktadır. Burada performans ölçütü olarak eğrinin altında kalan alan (AUC) yöntemi de “Şekil-10”da gösterilmiştir. Buradaki grafikte görülebileceği üzere 0.94 AUC değeriyle yüksek performanslı bir sınıflandırma gerçekleşmektedir. Sınıflandırma işleminin, veri elde etme yöntemlerinden biri olan yalnızca jiroskop verileriyle yapılması durumunda “Tablo-3” ve “Şekil-7”de görülebileceği üzere %90.6 doğrulukta sınıflandırma başarıları elde edilmiştir. Burada performans ölçütü olarak eğrinin altında kalan alan (AUC) yöntemi de “Şekil-11”de gösterilmiştir. AUC yöntemiyle yapılan değerlendirmede 0.96 AUC değeriyle yine yüksek performanslı bir sınıflandırma gerçekleştirilmektedir. Elde edilen bulgular literatürle karşılaştırıldığında genel olarak literatürde bulunan sensör verisi ile hareket tanımlama çalışmalarından elde edilen başarı oranları ile uyumlu ve hatta daha yüksek başarı oranlı sonuçlar elde edildiği ortaya konulmuştur.

#### 4. Sonuç

Bu çalışmada yapılan işlemler “Şekil-1”de belirtilen akış şemasında belirtilen şartlar dahilinde gerçekleştirilmiş ve sınıflandırma başarıları en yüksek model olan DVM modeli literatürle uyumlu bir şekilde %90.6 ile sınıflandırmayı gerçekleştirmiştir. Diğer başarılı modeller sırasıyla %89.9 başarı oranıyla Ensemble (Bagged Trees) ve %89.6 başarı oranıyla k-NN modeli olarak sonuçlanmıştır.

Burada karışıklık matrislerinden elde edilen sonuçlar doğrultusunda 1 numaralı hareket olan “Merdiven İnme” ve 4 numaralı hareket olan Merdiven Çıkma” verisi benzerliklerinden dolayı birbirlerine karışmaktadır. Bunun için gelecek çalışmalarda literatür taraması yapılarak tespit edilen farklı öznitelikler kullanılarak sınıflandırma başarıları artırılabilir.

Bu çalışmada kullanılan modellerde 3 numaralı hareket olan “Ayakta Durma” verisi tüm algoritmalarda yüksek doğrulukla ayrışabildiği görülmüştür.

Elde edilen bulguların genel olarak literatürle uyumlu ve literatüre katkı sağlayacak nitelikte olduğu görülmüştür.

Bu çalışmada elimizdeki veri setine bağlı olarak merdiven inme, merdiven çıkma, koşma ve ayakta durma sınıflandırmaları yapılmıştır. Ancak günlük hayatta oturma, yürüme, el yıkama, yüzme, zıplama vb. birçok hareket tipi bulunmaktadır. Çalışmanın geliştirilmesi için farklı veri setleri kullanılmalı veya bu aktiviteleri de içeren veri seti oluşturularak kullanılan modelin başarısı kontrol edilmelidir.

## Kaynakça

- [1] Malekzadeh, M., Clegg, R.G., Cavallaro, A., Haddadi, H., "Mobile sensor data anonymization", *In Proceedings of the International Conference on Internet of Things Design and Implementation (IoTDI '19)*, ACM, New York, NY, USA, 49-58, 2019.
- [2] Yan, W., Sun, Cang S., Yu, H., "A survey on wearable sensor modality centred human activity recognition in health care ", *Expert Systems With Applications*, 137 (2019) 167–190, 2019.
- [3] Jia, H., Chen, S., "Integrated data and knowledge driven methodology for human activity recognition", *Information Sciences*, 536(2020), 409-430, 2020.
- [4] Nweke, H.F., Teh, Y.W., Al-garadi, M.A., Alo, U.R., "Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges", *Expert Systems With Applications* 105 (2018), 233-261, 2020.
- [5] Subasi, A., Fllatah, A., Alzobidi, K., Brahim, T., Sarirete, A., "Smartphone-Based Human Activity Recognition Using Bagging and Boosting", *Procedia Computer Science* 163 (2019), 54-61, 2019.
- [6] Jobanputra, C., Bavishi, J., Doshi, N., "Human Activity Recognition: A Survey", *Procedia Computer Science* 155 (2019), 698-703, 2019.