

**BAŞKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI  
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS  
PROGRAMI**

**MAKİNE ÖĞRENMESİ İLE KIRILGANLIK TAHMİNİ VE ZAMAN  
ÖLÇEĞİ ETKİSİ**

**HAZIRLAYAN**

**MEHMET HALİT ÖZALP**

**YÜKSEK LİSANS TEZİ**

**ANKARA – 2024**



**BAŞKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI  
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS  
PROGRAMI**

**MAKİNE ÖĞRENMESİ İLE KIRILGANLIK TAHMİNİ VE ZAMAN  
ÖLÇEĞİ ETKİSİ**

**HAZIRLAYAN**

**MEHMET HALİT ÖZALP**

**YÜKSEK LİSANS TEZİ**

**TEZ DANIŞMANI**

**DR. ÖĞR. ÜYESİ DİDEM ÖLÇER**

**ANKARA – 2024**

**BAŞKENT ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı çerçevesinde Mehmet Halit ÖZALP tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 22 / 11 / 2024

**Tez Adı:** Makine Öğrenmesi ile Kırılma Tahmini ve Zaman Ölçeği Etkisi

**Tez jüri üyeleri (Unvanı, Adı-Soyadı, Kurumu)**

**İMZA**

Dr. Öğr. Üyesi İclal ÇETİN TAŞ, Ostim Teknik Üniversitesi .....

Dr. Öğr. Üyesi Didem ÖLÇER, Başkent Üniversitesi .....

Dr. Öğr. Üyesi Çağatay Berke ERDAŞ, Başkent Üniversitesi .....

**ONAY**

Prof. Dr. Dilek ÇÖKELİLER SERDAROĞLU

Fen Bilimleri Enstitüsü Müdürü

Tarih : ... / ... / .....

**BAŞKENT ÜNİVERSİTESİ**  
**FEN BİLİMLER ENSTİTÜSÜ**  
**YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU**

Tarih: 28 /11/ 2024

Öğrencinin Adı, Soyadı : Mehmet Halit Özalp

Öğrencinin Numarası : 22110033

Anabilim Dalı : Bilgisayar Mühendisliği

Programı : Tezli Yüksek Lisans

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Didem ÖLÇER

Tez Başlığı : Makine Öğrenmesi ile Kırılgnlık Tahmini ve Zaman Ölçeği Etkisi

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 68 sayfalık kısmına ilişkin, 28 /11/ 2024 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 2'dir. Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimededen daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası: .....

ONAY

Tarih: 28 /11/ 2024

Öğrenci Danışmanı Unvan, Adı, Soyadı, İmza:

.....

## TEŐEKKÜR

Bu yüksek lisans tezi sürecinde, bilgi ve deneyimleriyle bana rehberlik eden, akademik vizyonu ve desteęiyle çalışmamın her aşamasında yolumu aydınlatan danışman hocam Sayın Dr. Öğr. Üyesi Didem Ölçer'e derin teşekkürlerimi sunarım. Kendisinin değerli katkıları, bu çalışmanın şekillenmesinde önemli bir rol oynamıştır.

Ayrıca, sürecin her anında yanımda olan, sabır ve anlayışlarıyla bana güç veren aileme, varlıklarıyla sağladıkları manevi destek ve teşvikleri için en içten teşekkürlerimi ifade etmek isterim.

Bu çalışmanın her aşamasında, sevgisi, desteęi ve anlayışıyla beni motive eden ve bu süreci benim için daha anlamlı kılan sevgili eşim Tuęba Güran Özalp'e de gönülden teşekkür ederim. Onun varlığı, bu yolculuęu benim için hem kolaylaştırmış hem de anlamlı kılmıştır.

Bu tezin tamamlanmasında emeęi geçen ve yanımda olan herkese teşekkürlerimi sunarım.

Saygı ve teşekkürlerimle,

Mehmet Halit Özalp

Ankara - 2024

## ÖZET

**Mehmet Halit ÖZALP**

### **MAKİNE ÖĞRENMESİ İLE KIRILGANLIK TAHMİNİ VE ZAMAN ÖLÇEĞİ ETKİSİ**

**Başkent Üniversitesi Fen Bilimleri Enstitüsü**

**Bilgisayar Mühendisliği Anabilim Dalı**

**2024**

Yaşlanan dünya nüfusuyla birlikte yaşlı bireylerde kırılmanın erken aşamada tespit edilmesi, sağlık hizmetlerinin etkinliğini artırmakta ve bireylerin yaşam kalitesinin korunmasında önemli bir rol oynamaktadır. Kırılma, fiziksel, bilişsel ve psikolojik işlevlerdeki azalmalarla karakterize edilen bir durum olup, bireylerin bağımsız yaşam kapasitesini sınırlamaktadır. Bu çalışma, Makine Öğrenmesi algoritmalarını kullanarak giyilebilir cihazlardan elde edilen biyofiziksel veriler ile yaşlı bireylerde kırılma seviyelerinin belirlenmesi ve tahmin edilmesi üzerine odaklanmaktadır. Literatürdeki çalışmalar daha çok klinik gözlem ve testlere dayanmakta; ancak, giyilebilir cihazlar sayesinde elde edilen veri akışı ile gerçek zamanlı ve daha kapsamlı analizler mümkün olmaktadır. Bu tez kapsamında, ivmeölçer ve EKG sensörlerinden elde edilen veriler XGBoost, KNN, RF, 1D-CNN, MLP ve Naive Bayes gibi algoritmalarla analiz edilerek, kırılma tahmini için en uygun model ve zaman ölçeği belirlenmeye çalışılmıştır. Çalışmanın araştırma soruları, biyofiziksel verilerle kırılmanın doğru tahmin edilip edilemeyeceği, en verimli zaman aralığı ve hangi algoritmanın en iyi performansı gösterdiği üzerinedir. Beş bölümden oluşan bu tezde, konuya dair literatür incelemesi, veri toplama ve analiz yöntemleri, algoritmaların uygulamaları ve araştırmanın sonuçları ele alınmaktadır.

**ANAHTAR KELİMELER:** Kırılma Sınıflandırma, Akselerometre, EKG, Makine Öğrenmesi, Öznitelik Çıkarımı.

## **ABSTRACT**

**Mehmet Halit ÖZALP**

### **FRAILITY ESTIMATION WITH MACHINE LEARNING AND TIME SCALE EFFECT**

**Başkent University Insitute of Science and Engineering**

**Department of Computer Engineering**

**2024**

With the aging world population, early detection of frailty in elderly individuals plays a crucial role in enhancing the effectiveness of healthcare services and preserving individuals' quality of life. Frailty is characterized by declines in physical, cognitive, and psychological functions, which limit an individual's capacity for independent living. This study focuses on using machine learning algorithms to determine and predict frailty levels in elderly individuals based on biophysical data obtained from wearable devices. While most studies in the literature rely on clinical observation and tests, the continuous data flow from wearable devices enables more comprehensive and real-time analyses. In this thesis, accelerometer and ECG sensor data will be analyzed using algorithms such as XGBoost, KNN, RF, 1D-CNN, MLP, and Naive Bayes to identify the most suitable model and time window for frailty prediction. The research questions of this study address whether frailty can be accurately predicted using biophysical data, what the most effective time window is, and which algorithm performs best. Structured in five main chapters, this thesis covers a review of the relevant literature, data collection and analysis methods, algorithm applications, and findings of the research.

**KEYWORDS:** Frailty Classification, Accelerometer, ECG, Machine Learning, Feature Extraction.

## İÇİNDEKİLER

|   |           |
|---|-----------|
| TEŞEKKÜR.....   | i         |
| ÖZET.....   | ii        |
| ABSTRACT .....  | iii       |
| İÇİNDEKİLER.....  | iv        |
| ŞEKİLLER LİSTESİ.....   | vi        |
| TABLolar LİSTESİ.....   | vii       |
| SİMGELER VE KISALTMALAR .....   | ix        |
| <b>1. GİRİŞ.....</b>  | <b>1</b>  |
| <b>1.1. Araştırma Konusunun Tanıtımı.....</b>                           | <b>1</b>  |
| <b>1.2. Alan Bilgisi.....</b>   | <b>1</b>  |
| <b>1.3. Problem Tanımı.....</b>   | <b>2</b>  |
| <b>1.4. Araştırmanın Amacı.....</b>                                     | <b>2</b>  |
| <b>1.5. Araştırmanın Kapsamı.....</b>                                   | <b>2</b>  |
| <b>1.6. Araştırma Soruları.....</b>                                     | <b>3</b>  |
| <b>1.7. Tezin Yapısının Özeti.....</b>                                  | <b>3</b>  |
| <b>2. ALAN BİLGİSİ VE LİTERATÜR.....</b>                                | <b>4</b>  |
| <b>3. YÖNTEM VE MATERYAL.....</b>                                       | <b>14</b> |
| <b>3.1. Veri Seti.....</b>  | <b>14</b> |
| <b>3.2. Veri Ön İşleme.....</b>   | <b>15</b> |
| <b>3.3. Kullanılan Algoritmalar.....</b>                                | <b>16</b> |
| <b>3.3.1. K-En Yakın Komşu (KNN) Algoritması.....</b>                   | <b>16</b> |
| <b>3.3.2. Rastgele Orman (RF) Algoritması.....</b>                      | <b>17</b> |
| <b>3.3.3. Ekstrem Gradyan Artırma (XGBoost) Algoritması.....</b>        | <b>18</b> |
| <b>3.3.4. Tek Boyutlu Evrişimli Sinir Ağı (1D-CNN) Algoritması.....</b> | <b>18</b> |
| <b>3.3.5. Naive Bayes Algoritması.....</b>                              | <b>19</b> |
| <b>3.3.6. Çok Katmanlı Algılayıcı (MLP) Algoritması.....</b>            | <b>20</b> |
| <b>3.4. Performans Değerlendirmesi İçin Kullanılan Metrikler.....</b>   | <b>20</b> |
| <b>3.4.1. Straifield K-Fold.....</b>                                    | <b>20</b> |
| <b>3.4.2. Doğruluk (Accuracy).....</b>                                  | <b>21</b> |
| <b>3.4.3. Kesinlik (Precision).....</b>                                 | <b>22</b> |
| <b>3.4.4. Duyarlılık (Recall).....</b>                                  | <b>22</b> |

|   |    |
|---|----|
| 3.4.5. F1 Skoru .....                           | 22 |
| 3.4.6. MCC (Matthews Korelasyon Katsayısı)..... | 23 |
| 3.5. Deneysel Kurulum .....                     | 23 |
| 3.6. Deneysel Sonular .....                    | 26 |
| 4. TARTIŐMA .....                               | 49 |
| 5. SONU .....                                  | 67 |
| 6. KAYNAKLAR.....                               | 69 |



## ŞEKİLLER LİSTESİ

|  | <b>Sayfa</b> |
|--|--------------|
| Şekil 3.1. K=4 için örnek KNN .....                      | 17           |
| Şekil 3.2. Rastgele Orman Sınıflandırma yapısı [38]..... | 18           |
| Şekil 3.3. Örnek 1D-CNN Modeli [41] .....                | 19           |
| Şekil 3.4. SKCV K=3 .....                                | 21           |
| Şekil 3.5. Veri Setleri Oluşturma Adımları .....         | 25           |



## TABLolar LİSTESİ

|  | <b>Sayfa</b> |
|--|--------------|
| Tablo 3.1. EFS skoruna göre Kırılgnlık sınıflandırmaları ..... | 24           |
| Tablo 3.2. İvmeölçer STAIR 2 saniye .....                      | 26           |
| Tablo 3.3. İvmeölçer STAIR 4 saniye .....                      | 26           |
| Tablo 3.4. İvmeölçer STAIR 8 saniye .....                      | 27           |
| Tablo 3.5. EKG STAIR 2 saniye .....                            | 27           |
| Tablo 3.6. EKG STAIR 4 saniye .....                            | 28           |
| Tablo 3.7. EKG STAIR 8 saniye .....                            | 28           |
| Tablo 3.8. İvmeölçer & EKG STAIR 2 saniye .....                | 29           |
| Tablo 3.9. İvmeölçer & EKG STAIR 4 saniye .....                | 29           |
| Tablo 3.10. İvmeölçer & EKG STAIR 8 saniye .....               | 30           |
| Tablo 3.11. İvmeölçer TUG 2 saniye.....                        | 30           |
| Tablo 3.12. İvmeölçer TUG 4 saniye.....                        | 31           |
| Tablo 3.13. İvmeölçer TUG 8 saniye.....                        | 31           |
| Tablo 3.14. EKG TUG 2 saniye.....                              | 32           |
| Tablo 3.15. EKG TUG 4 saniye.....                              | 32           |
| Tablo 3.16. EKG TUG 8 saniye.....                              | 33           |
| Tablo 3.17. İvmeölçer & EKG TUG 2 saniye .....                 | 33           |
| Tablo 3.18. İvmeölçer & EKG TUG 4 saniye .....                 | 34           |
| Tablo 3.19. İvmeölçer & EKG TUG 8 saniye .....                 | 34           |
| Tablo 3.20. İvmeölçer Veloergometri 2 saniye .....             | 35           |
| Tablo 3.21. İvmeölçer Veloergometri 4 saniye .....             | 35           |
| Tablo 3.22. İvmeölçer Veloergometri 8 saniye .....             | 36           |
| Tablo 3.23. EKG Veloergometri 2 saniye .....                   | 36           |
| Tablo 3.24. EKG Veloergometri 4 saniye .....                   | 37           |

|   |    |
|---|----|
| Tablo 3.25. EKG Veloergometri 8 saniye .....            | 37 |
| Tablo 3.26. İvmeölçer & EKG Veloergometri 2 saniye..... | 38 |
| Tablo 3.27. İvmeölçer & EKG Veloergometri 4 saniye..... | 38 |
| Tablo 3.28. İvmeölçer & EKG Veloergometri 8 saniye..... | 39 |
| Tablo 3.29. İvmeölçer GAINT 2 saniye .....              | 39 |
| Tablo 3.30. İvmeölçer GAINT 4 saniye .....              | 40 |
| Tablo 3.31. İvmeölçer GAINT 8 saniye .....              | 40 |
| Tablo 3.32. EKG GAINT 2 saniye .....                    | 41 |
| Tablo 3.33. EKG GAINT 4 saniye .....                    | 41 |
| Tablo 3.34. EKG GAINT 8 saniye .....                    | 42 |
| Tablo 3.35. İvmeölçer ve EKG GAINT 2 saniye .....       | 42 |
| Tablo 3.36. İvmeölçer ve EKG GAINT 4 saniye .....       | 43 |
| Tablo 3.37. İvmeölçer ve EKG GAINT 8 saniye .....       | 43 |
| Tablo 3.38. İvmeölçer 6MWT 2 saniye .....               | 44 |
| Tablo 3.39. İvmeölçer 6MWT 4 saniye .....               | 44 |
| Tablo 3.40. İvmeölçer 6MWT 8 saniye.....                | 45 |
| Tablo 3.41. EKG 6MWT 2 saniye .....                     | 45 |
| Tablo 3.42. EKG 6MWT 4 saniye .....                     | 46 |
| Tablo 3.43. EKG 6MWT 8 saniye .....                     | 46 |
| Tablo 3.44. İvmeölçer & EKG 6MWT 2 saniye.....          | 47 |
| Tablo 3.45. İvmeölçer & EKG 6MWT 4 saniye.....          | 47 |
| Tablo 3.46. İvmeölçer & EKG 6MWT 8 saniye.....          | 48 |
| Tablo 4.1. Özet Sonuçlar .....                          | 65 |

## SİMGELER VE KISALTMALAR

|          |  |
|----------|--|
| ML       | Machine Learning (Makine Öğrenmesi)                                  |
| KNN      | K-Nearest Neighbors (K-En Yakın Komşu)                               |
| RF       | Random Forest (Rastgele Orman)                                       |
| XGBoost  | Extreme Gradient Boosting (Ekstrem Gradyan Artırma)                  |
| CNN      | Convolutional Neural Network (Evrşimsel Sinir Ağı)                   |
| MLP      | Multi-Layer Perceptron (Çok Katmanlı Algılayıcı)                     |
| EKG      | Elektrokardiyografi  |
| MCC      | Matthews Correlation Coefficient (Matthews Korelasyon Katsayısı)     |
| ROC      | Receiver Operating Characteristic (Alıcı İşletim Karakteristiği)     |
| SKCV     | Stratified K-Fold Cross-Validation (Katmanlı K-Kat Çapraz Doğrulama) |
| TUG      | Timed Up and Go (Zamanlı Kalk ve Git)                                |
| 6MWT     | 6-Minute Walk Test (6 Dakika Yürüme Testi)                           |
| EFS      | Edmonton Frail Scale   |
| GPS      | Global Positioning System (Küresel Konumlandırma Sistemi)            |
| SVM      | Support Vector Machines (Destek Vektör Makineleri)                   |
| LR       | Logistic Regression (Lojistik Regresyon)                             |
| LightGBM | Light Gradient Boosting Machine (Hafif Gradyan Artırma Makinesi)     |

# 1. GİRİŞ

## 1.1. Araştırma Konusunun Tanıtımı

Yaşlanan dünya nüfusu ile yaşlı bireylerdeki kırılabilirlik durumu giderek daha önemli bir sağlık sorunu haline gelmiştir. Kırılabilirlik, genellikle fiziksel, bilişsel ve psikolojik işlevlerde görülen azalmalarla karakterize edilen ve bireyin bağımsız yaşam kapasitesini sınırlayan bir durumdur [1]. Yaşlı bireylerde kırılabilirliğin erken aşamada tespit edilmesi, sağlık hizmetlerinin etkinliğini artırmakta ve bireyin yaşam kalitesinin korunmasında kritik bir rol oynamaktadır [2]. Bu bağlamda, teknoloji ve yapay zeka tabanlı çözümlerle kırılabilirliğin belirlenmesi ve tahmin edilmesi büyük bir önem kazanmıştır.

Bu tez, Makine Öğrenmesi (ML) algoritmalarını kullanarak yaşlı bireylerde kırılabilirlik durumunun belirlenmesi ve kırılabilirlik seviyesinin tahmin edilmesine odaklanmaktadır. Giyilebilir cihazlardan elde edilen biyofiziksel verilerin analizi sayesinde, yaşlı bireylerin sağlık durumlarını sürekli olarak izlemek ve kırılabilirlik seviyelerini doğru bir şekilde tahmin etmek mümkün hale gelmiştir.

## 1.2. Alan Bilgisi

Kırılabilirlik tespiti ve tahmini üzerine yapılan çalışmalar, ağırlıklı olarak klinik gözlem ve testlere dayanmaktadır. Ancak, giyilebilir cihazlar sayesinde elde edilen sürekli biyofiziksel veri akışı, daha gerçek zamanlı ve kapsamlı bir yaklaşım sunmaktadır. Literatürde, ML algoritmalarının bu veriler üzerinde uygulanmasıyla kırılabilirlik tahminine yönelik çeşitli çalışmalar mevcuttur. Kırılabilirliğin tahmini özelinde ivmeölçer, EKG gibi sensörlerden alınan veriler kullanılarak, yürüyüş analizi, denge durumu gibi parametreler üzerinden tahminler yapılmakta ancak zaman ölçeklerinin etkisi sorgulanmamaktadır. Bununla birlikte, bu çalışmaların büyük kısmı sınırlı veri setleri ile yapılmış olup, genel geçer modellerin geliştirilmesi için daha kapsamlı araştırmalara ihtiyaç duyulmaktadır. Bu tez çalışması, literatürdeki boşluğu doldurmak amacıyla daha geniş bir veri seti ve çeşitli algoritmaların ve zaman ölçeklerinin karşılaştırmalı analiziyle kırılabilirlik tespiti ve tahminine yönelik bir model geliştirmeyi hedeflemektedir.

### **1.3. Problem Tanımı**

Bu çalışmada ele alınan temel problem, yaşlı bireylerin kırılgnlık seviyelerinin, giyilebilir cihazlardan elde edilen biyofiziksel veriler kullanılarak doğru bir şekilde tahmin edilip edilemeyeceğidir. Geleneksel klinik yöntemler, zaman alıcı ve sınırlı veri sağlayan yöntemlerdir. Ancak, giyilebilir cihazlar yardımıyla bireylerden günlük yaşam aktiviteleri sırasında sürekli ve kesintisiz veri toplanabilmektedir. Bu verilerin doğru analiz edilmesi, yaşlı bireylerde kırılgnlık durumunun daha erken ve etkili bir şekilde belirlenmesine katkı sağlamıştır. Çalışmada kullanılan veriler, ivmeölçer ve EKG sensörlerinden alınmış olup, bu veriler üzerinden çeşitli ML algoritmalarıyla kırılgnlık seviyeleri tahmin edilmiştir.

### **1.4. Araştırmanın Amacı**

Bu tezin amacı, yaşlı bireylerde kırılgnlık seviyesini ML algoritmaları yardımıyla tahmin edebilecek bir model geliştirmektir. Çalışmada Polar H10 gibi giyilebilir cihazlardan elde edilen biyofiziksel veriler kullanılmaktadır. Bu veriler ivmeölçer ve EKG sinyallerini içermekte olup, kırılgnlık tespiti ve tahmini için çeşitli algoritmalar üzerinde analiz edilmektedir. Ekstrem Gradyan Arttırma (XGBoost), K En Yakın Komşu (KNN), Rastgele Orman (RF), Çok Katmanlı Algılayıcı (MLP) ve 1D-CNN (Tek Boyutlu Evrişimli Sinir Ağı) gibi ML algoritmaları, bu veriler üzerinde uygulanmış ve en başarılı modelin belirlenmesi hedeflenmektedir. Araştırmanın nihai hedefi, en uygun algoritma ve zaman ölçeği kullanılarak kırılgnlık seviyelerinin doğru bir şekilde tahmin edilmesini sağlamaktır.

### **1.5. Araştırmanın Kapsamı**

Bu araştırma, giyilebilir cihazlardan elde edilen biyofiziksel verilerin analizi ve ML algoritmalarının bu veriler üzerinde uygulanmasıyla kırılgnlık seviyesinin tahmin edilmesi üzerine odaklanmaktadır. Çalışmada kullanılan veriler, EKG ve ivmeölçer sinyallerini içermektedir. Veriler, farklı zaman ölçeklerine bölünerek analiz edilmiş ve bu pencereler üzerinden çeşitli ML algoritmalarıyla sonuçlar karşılaştırılmıştır. Çalışma kapsamında sadece belirli algoritmalar (XGBoost, KNN, RF, 1D-CNN, MLP ve Naive Bayes) ele alınmış ve diğer potansiyel algoritmalar dışarıda bırakılmıştır.

## 1.6. Arařtırma Soruları

Bu tez alıřmasında yanıtlanması gereken temel arařtırma soruları řunlardır:

1. Giyilebilir cihazlardan elde edilen biyofiziksel veriler kullanılarak, yařlı bireylerde kırılgnlık seviyesi doęru bir řekilde tahmin edilebilir mi?
2. Kırılgnlık seviyesinin tahmini iin en uygun zaman leęi nedir?
3. Hangi ML algoritması, biyofiziksel veriler zerinde en iyi performansı gstermektedir?

## 1.7. Tezin Yapısının zeti

Bu tez, beř ana blmden oluřmaktadır. İlk blmde, yařlı bireylerde kırılgnlık kavramı ve kırılgnlıęın neden nemli olduęu zerinde durulmaktadır. İkinci blmde, literatrde bu konuya dair yapılan alıřmalar incelenerek ve mevcut bořluklar ele alınmaktadır. nc blm, alıřmada kullanılan veri seti ve veri toplama yntemlerini detaylandırılmaktadır. Drdnc blmde, ML algoritmalarının uygulanıřı ve sonular verilmektedir. Beřinci ve son blmde ise arařtırmanın sonuları deęerlendirilip, bu sonuların gelecekte yapılacak alıřmalar iin neriler sunulmaktadır.

## 2. ALAN BİLGİSİ VE LİTERATÜR

Kırılgnlık, ilerleyen yaş ile ortaya çıkan ve oldukça ciddi sağlık sorunlarına sebebiyet veren geriatric bir sendromdur. Bu nedenle, incelenmesi ve tedavi edilmesi, yaşlı bireylerin daha sağlıklı ve kaliteli bir hayat sürmesi için oldukça önem arz eder. Kırılgnlığın tespiti için çeşitli yöntem arayışlar devam etmektedir. Bu bölümde literatürde bu alanlarda yapılmış çalışmalara yer verilmiştir.

Rolfson ve arkadaşları [3], kırılgnlık tespiti için yöntemlerden biri olan Edmonton Frail Scale (EFS) aracının güvenilirliğini denemişlerdir. EFS'nin klinik ortamlarda kullanımını kolaylaştıracak kısa ve kullanıcı dostu bir tarama aracı olarak geçerliliğini araştırmayı amaçlamaktadır. EFS kısaca yaşlı bireylerde kırılgnlık durumunu değerlendirmek için kullanılan bir tarama aracıdır. EFS, 10 farklı alanı kapsayan bir ölçek olarak tasarlanmıştır ve kırılgnlığın çok boyutlu (10 farklı alan) yapısını değerlendirmeyi amaçlamaktadır. Araştırmada, 65 yaş ve üzeri 158 katılımcı yer almıştır. Bu katılımcılar, Comprehensive Geriatric Assessment (CGA) değerlendirmesi için başvurmuşlardır. EFS, eğitim almamış bir araştırma asistanı tarafından uygulanmış ve sonuçlar, CGA ile karşılaştırılmış ve EFS'nin güvenilir bir araç olduğunu göstermişlerdir.

Bir başka çalışmada Fried ve arkadaşları [4], kırılgnlığın fiziksel, biyolojik ve klinik göstergelerini belirlemeyi ve bu sendromun yaşlı nüfus içindeki yaygınlığını ve etkilerini değerlendirmeyi hedeflemektedir. 65 yaş ve üzerindeki 5317 katılımcının dahil olduğu Kardiyovasküler Sağlık Çalışmasından elde edilen verilere dayanmaktadır. Kırılgnlık fenotipinin geliştirilmesi için, istemsiz kilo kaybı, kendini yorgun hissetme, kas zayıflığı (el kavrama gücü), yavaş yürüme hızı ve düşük fiziksel aktivite düzeyi gibi beş klinik kriter kullanılmıştır. Bu kriterlerden en az üçünün varlığı, bireyin kırılgn olarak sınıflandırılması için yeterli sayılmıştır. Kırılgnlığın tespiti için standart bir fenotipin kullanılabileceği sonucuna ulaşmıştır.

Kırılgnlık tespiti için kullanılan yöntemlerin haricinde kırılgnlıkla ilgili yapılan çalışmalar incelendiğinde 3 ana başlık öne çıkmaktadır. Bu konuların her biri kırılgnlık konusunda oldukça önemli bir yer tutmaktadır. İlk olarak kırılgnlığın tespitinde de kullanılan ve kırılgnlık gelişiminde en önemli faktör olarak gösterilen Fiziksel Aktivite ile kırılgnlığın arasındaki ilişkiyi inceleyen çalışmalardır. Bir diğeri, giyilebilir cihazların doğrulanması ve bu cihazlardan toplanan veriler ile kırılgnlığın tespit edilmesini inceleyen

çalışmalardır. Son olarak ilgili tez içinde önem arz eden testler esnasında toplanan verilerin makine öğrenmesi (ML) algoritmaları kullanılarak işlenmesi ve kırılğanlığın tahmin edilmesi ile ilgili yapılan çalışmalardır. Bahsi geçen bu 3 grupta literatürdeki çalışmalar bu bölümde yer almaktadır.

Fiziksel Aktivite ile kırılğanlığın yadsınamaz bir ilişkisi olduğu yargısı literatüre oldukça hakimdir. Bu nedenle genel olarak kırılğanlık tespiti için kullanılan testler de Fiziksel Aktivite tabanlıdır. Örneğin Martínez ve arkadaşları [5], özellikle denge bozukluklarının tespiti için kullanılan hızlanma ve yönelim sinyallerinin frekans içeriklerinin kırılğanlık durumu ile ilişkisini araştırmışlardır. Yaşları 40 ile 80 arasında değişen 14'ü kırılğan, 18'i ön-kırılğan ve 24 sağlıklı toplam 56 gönüllü katılmıştır. Katılımcıların denge performansları, üç eksenli atalet manyetik sensör kullanılarak ölçülmüştür. Sessiz duruş denge testleri sırasında, sensörler tarafından elde edilen hızlanma ve yönelim sinyalleri incelenmiştir. Farklı denge pozisyonları (ayaklar bitişik ve yarı tandem pozisyonunda, gözler açık ve kapalı) altında ölçümler alınmıştır. Kırılğan bireylerde hızlı denge düzeltme yeteneğinin azaldığını ve yüksek frekans bileşenlerinin kırılğanlığı ayırt etmede faydalı olabileceğini göstermektedir.

Millor ve arkadaşları [6], yaşlı bireylerde kırılğanlığın tespitini geliştirmek için 30 saniyelik sandalye kalkma testi sırasında bir tekil atalet ünitesi (IU) kullanarak elde edilen kinematik parametrelerin kırılğanlık seviyelerini (kırılğan, ön-kırılğan, sağlıklı) ayırt etmede ne kadar etkili olduğunu incelemişlerdir. Yaşları ortalama 54 olan 18 sağlıklı, yaşları ortalama 78 olan 16 ön- kırılğan ve yaşları ortalama 85 olan 13 kırılğan katılımcı, 30 saniyelik sandalye kalkma testini gerçekleştirirken, gövde hareketleri bir atalet ünitesi ile ölçülmüş ve kinematik parametreler farklı kırılğanlık seviyelerine göre karşılaştırılmıştır. Ayağa kalkma ve oturma aşamalarındaki hız zirveleri ve "modifiye impuls" parametreleri, farklı kırılğanlık seviyelerine sahip bireyleri açık bir şekilde ayırt etmiştir.

Galán-Mercant ve arkadaşları [7], 10 metre Genişletilmiş Kalkma ve Yürüme Testi (ETGUG) sırasında bir akıllı telefonun (iPhone 4) ivmeölçer ve jiroskopunu kullanarak bu farkları ölçüp değerlendirmeyi amaçlamışlardır. 65 yaş üstü 30 katılımcı (14 kırılğan, 16 sağlıklı) ile İvme ve açısal hız ölçümleri iPhone 4 cihazı kullanılarak gövde üzerinde yapılmıştır. Elde edilen veriler kırılğan ve sağlıklı bireyler arasındaki farklılıkları değerlendirmek için istatistiksel olarak analiz edilmiştir. Kırılğan bireylerde, gövde kinematiği açısından sağlıklı bireylere kıyasla önemli farklılıklar gözlenmiştir. Kırılğan

grubunda ivme ve açısal hız değişkenlerinin daha düşük olduğu ve bu bireylerin dönüş geçişlerini daha yavaş ve dikkatli bir şekilde gerçekleştirdiği tespit edilmiştir. Aynı yıl başka bir çalışmada Galán-Mercant ve arkadaşları [8], kırılğan ve sağlıklı yaşlı bireylerde oturma-kalkma ve kalkma-oturma geçişleri sırasında gövde kinematiği arasındaki farkları incelemek için iPhone 4'ün dahilindeki mobil atalet sensörünü kullanmışlardır. Yaşları 65'in üzerinde olan 14'ü kırılğan, 16'sı sağlıklı 30 katılımcıyı, Fried kriterlerine göre kırılğanlık durumu açısından sınıflandırılmıştır. iPhone 4'ün ivmeölçer sensörleri kullanılarak gövde üzerinde ölçümler yapılmıştır. Katılımcılar, standart bir sandalyede üç kez ardışık oturma-kalkma ve kalkma-oturma geçişi gerçekleştirmiştir. Kırılğan grupta ivme ve açısal hız değişkenleri daha düşük seviyelerde gözlenmiş ve bu bireylerin testlerdeki geçişlerini daha dikkatli ve kısıtlı bir şekilde gerçekleştirdikleri tespit edilmiştir.

Greene ve arkadaşları [9], yaşlı yetişkinlerde kırılğanlık ve düşme geçmişini sınıflandırmak için, eylemsizlik ve basınç sensörleri kullanarak denge ve hareketlilik verilerini elde edip bu verilerle kırılğanlık ve düşme riski tahmin modelleri geliştirmeyi hedeflemiştir. Ortalama yaşı  $75.9 \pm 6.6$  yıl olan 124 yaşlı birey Fried kırılğanlık kriterlerine göre kırılğan, ön- kırılğan ve kırılğan olmayan olarak sınıflandırılmıştır. Timed Up and Go (Zamanlı Kalk ve Git – TUG), Five Times Sit to Stand (Beş Defa Otur-Kalk Testi – FTSS) ve quiet standing (sessiz duruş) denge testleri esnasında elde edilen sensör verileri Support Vector Machines (SVM) tabanlı sınıflandırma modelleri kullanılarak incelenmiştir. Sonuçlar kırılğanlık ve düşme geçmişini sınıflandırmada büyük oranda doğru sonuçlar verdiği göz önünde bulundurularak sensör tabanlı değerlendirmelerin kırılğanlık ve düşme riski belirlemede etkili olabileceğini göstermektedir. Aynı yıl bir başka çalışmada Greene ve arkadaşları [10], TUG testi esnasında atalet sensörleri kullanarak kırılğanlığı hızlı ve otomatik bir şekilde tespitinin yapılıp yapılamayacağını değerlendirmiştir. 399 katılımcı ile yapılan çalışmada toplanılan veriler geliştirilen regresyon modeli ile analiz edilmiştir. Yapılan çalışmanın doğru tahmin oranı %70'in üzerinde çıkmıştır.

Jansen ve arkadaşları [11], kırılğan ve kırılğan olmayan yaşlı bireylerde fiziksel aktivite seviyelerini objektif olarak karşılaştırmak için iç mekan ve dış mekan aktivitelerini incelemiş ve fiziksel aktivite farklarını belirlemeye çalışmıştır. 65-89 yaş aralığındaki 74'ü kırılğan olmayan, 10'u kırılğan toplam 84 yaşlı birey üzerinde gerçekleştirilmiştir. Fiziksel aktivite yedi gün boyunca İvmeölçer ve GPS cihazları kullanılarak ölçülmüştür. Kırılğan

olmayan yaşlılar, orta-yoğun fiziksel aktiviteleri ev dışında daha fazla gerçekleştirirken, kırılğan olan yaşlılar daha çok evde hareketsiz vakit geçirmiştir.

Chen ve arkadaşları [12], fiziksel aktivite düzeyinin kırılğanlık fenotipini belirlemede kullanılabilirliğini araştırmıştır. 65 yaş ve üzeri 1527 Japon katılımcı, bir ivmeölçer ile kullanılarak fiziksel aktiviteleri objektif olarak ölçülmüştür. Veriler çoklu adımlı lojistik regresyon modellerinde kullanılmıştır. Yaş, zayıf sağlık algısı, depresif ve anksiyete semptomları, alkol tüketmemek, sosyal aktivitelere katılmamak ve bilişsel bozukluk gibi faktörler, kırılğanlık ile güçlü bir şekilde ilişkili bulunmuştur.

Castaneda-Gameros ve arkadaşları [13], farklı etnik ve dini gruplara mensup yaşlı göçmen kadınlarda fiziksel aktivite, hareketsiz zaman ve kırılğanlık arasındaki ilişkiyi incelemektedir. 60 yaş ve üzerinde, 60 göçmen kadın fiziksel aktivite ve hareketsiz zaman seviyeleri, İvmeölçer kullanılarak objektif olarak ölçülmüştür. Veriler ışığında yapılan çalışmaya göre, kırılğanlık durumuna bakılmaksızın katılımcıların oldukça hareketsiz olduklarını göstermiştir.

Huisingh-Scheetz ve arkadaşları [14], kırılğanlık durumunun ve diğer demografik ve sağlık özelliklerinin yaşlı bireylerdeki fiziksel aktivite düzeyleri üzerindeki etkisini belirlemeyi amaçlamıştır. 651 yaşlı bireyden toplanan saatlik İvmeölçer verileri analiz etmiş ve doğrusal regresyon modelleri kullanılarak, saatlik fiziksel aktivite düzeylerinin kırılğanlık skorları ve diğer demografik değişkenlerle ilişkisi incelenmiştir. Kırılğanlık skorları daha düşük aktivite düzeyleri ile ilişkili bulunmuştur ve her bir kırılğanlık puanı, saatlik ortalama aktivite düzeyinde %7'lik bir azalma ile ilişkilendirilmiştir.

Yuki ve arkadaşları [15], Japonya'da toplulukta yaşayan yaşlı bireylerde günlük fiziksel aktivitenin kırılğanlık gelişimini nasıl etkilediğini incelemiştir. 401 yaşlı bireyin fiziksel aktiviteleri, tek eksenli İvmeölçer ile kaydedilmiştir. Kırılğanlık, Fried kırılğanlık kriterlerine göre değerlendirilmiş ve katılımcılar 3 grupta sınıflandırılmıştır. Kırılğanlık gelişimini öngörmek için genel tahmin denklemleri kullanılarak analiz edilmiştir. Fiziksel aktivitenin az olduğu bireylerde kırılğanlık gelişimi riskinin arttığı görülmüştür.

Jansen ve arkadaşları [16], kırılğanlık durumunun, laboratuvar ortamında değerlendirilen motor kapasite ile günlük hayatta gözlemlenen mobilite performansı arasındaki ilişkiyi moderatör olarak nasıl etkilediğini araştırmıştır. 40'ı sağlıklı, 53'ü ön-

kırılğan ve 19'u kırılğan olmak üzere toplam 112 yaşlı yetişkin üzerinde, motor kapasitesi ve Mobilite performansı ölçülmüştür. İstatistiksel analizler, hiyerarşik doğrusal regresyon modelleri kullanılarak yapılmış ve kırılğanlık durumunun bu ilişkiye olan etkisi araştırılmıştır. Kırılğan olmayan bireylerde motor kapasite ile mobilite performansı arasında anlamlı bir ilişki olmadığını, ancak ön-kırılğan ve kırılğan bireylerde bu ilişkinin belirgin olduğunu göstermiştir.

Ziller ve arkadaşları [17], fiziksel aktivite kriterlerinin kırılğanlık fenotipinin belirlenmesinde önemli bir etkiye sahip olup olmadığını ve bu farklılıkların prevalans oranlarını nasıl etkilediğini belirlemeye çalışmışlardır. 65 yaş ve üzerindeki 47 katılımcı üzerinde çapraz kesitsel bir analiz yapılmıştır. Kırılğanlık durumu, kırılğanlık fenotipi kriterlerine göre değerlendirilmiş, düşük fiziksel aktivite kriteri ise beş farklı anket ve bir İvmeölçer kullanılarak belirlenmiştir. Fiziksel aktivite değerlendirme yöntemi ve kesme noktası seçimlerinin kırılğanlık prevalansını önemli ölçüde etkilediğini göstermektedir. Standartlaştırılmış fiziksel aktivite kriterleri oluşturulmadıkça, farklı çalışmalarda yayımlanan prevalans oranlarının karşılaştırılmasının zor olacağı görülmüştür.

Kikuchi ve arkadaşları [18], Japonya'da yaşayan yaşlı bireyler arasında fiziksel aktivite yoğunluğu ve belirli süreli hareketsiz zamanın kırılğanlık durumu ile olan ilişkisini incelemiştir. 65-84 arasında 511 yaşayan yaşlı bireyin fiziksel aktiviteleri ve hareketsiz zamanları, üç eksenli bir İvmeölçer kullanılarak ölçülmüş ve istatistiksel yöntemlerle analiz edilmiştir. Sağlıklı bireylerde fiziksel aktivitenin daha yoğun olduğu kırılğan veya ön-kırılğan bireylerde daha az yoğunlukta fiziksel aktivite tespit edilmiştir.

Fiziksel aktivite ile kırılğanlık arasındaki bağlantıyı inceleyen çalışmalarda fiziksel aktivitenin kırılğanlık üzerinde büyük bir etkisi olduğu, din, dil, ırk fark etmeksizin tüm çalışmalarda aynı sonuçlara ulaşıldığı görülmektedir. Ancak ölçüm yapılan cihazlar yani sensörlerin güvenilir mi olduğu sorusu önem arz etmektedir. Bu nedenle kırılğanlık ile ilgili bir diğer önemli çalışma alanı ise ölçüm yapılan sensörlerin güvenilirliği ve sağlıklı bir ölçüm yapıp yapmadığıdır.

Theou ve arkadaşları [19], kırılğanlık seviyeleri arasında fiziksel aktiviteyi ölçmek için kullanılan araçların karşılaştırmasını yapmak ve hangi fiziksel aktivite ölçüm araçlarının kırılğanlık ile en yakından ilişkili olduğunu belirlemeye çalışmıştır. 63-90 yaş aralığındaki 50 toplulukta yaşayan kadın, fiziksel aktiviteler, 10 saat boyunca objektif olarak bir

İvmeölçer, kalp atış hızı monitörü, taşınabilir elektromiyografi cihazı ve GPS kullanılarak ölçülmüştür. İvmeölçer adım sayısı ve kırılğanlığın en güçlü korelasyona sahip olduğu ayrıca İvmeölçer, kalp atış hızı monitörü, taşınabilir elektromiyografi cihazı ve GPS verilerinin birlikte kullanımının doğruluk oranını arttırdığı görülmüştür.

Schwenk ve arkadaşları [20], yaşlı bireylerde yürüme, denge ve fiziksel aktiviteyi değerlendirmek için kullanılan giyilebilir sensör tabanlı ev içi değerlendirmelerin, kırılğanlık seviyelerini (kırılğan olmayan, ön-kırılğan, kırılğan) ayırt etme yeteneğini incelemiştir. 44'ü sağlıklı 60'ı ön-kırılğan ve 21'i kırılğan olan toplam 125 yaşlı bireyin yürüme hızları, adım uzunluğu, çift destek süresi, denge ve fiziksel aktivite düzeyleri gibi parametreler, kişilerin evlerinde giyilebilir sensörler kullanılarak ölçülmüştür. Veri analizi, lojistik regresyon ve çok değişkenli lojistik regresyon modelleri kullanılarak gerçekleştirilmiştir. Yürüme hızı, kalça salınımı ve günlük adım sayısı, ön-kırılğan katılımcıların tespitinde en hassas parametreler olarak belirlenmiştir.

Toosizadeh ve arkadaşları [21], kırılğanlık durumunu belirlemek için yürüme analizine dayanmayan, yenilikçi bir üst ekstremite hareket değerlendirme yöntemi geliştirmeye ve bu yöntemi kablosuz sensörlerle doğrulamaya çalışmıştır. 65 yaş ve üzeri 117 toplulukta yaşayan yaşlı bireyler üzerinde gerçekleştirilmiştir. Katılımcılara, üst kola ve ön kola kablosuz sensörler takılarak, her iki kolda 20 saniyelik dirsek fleksiyonu egzersizi yaptırılmıştır. Veriler, Fried kırılğanlık kategorileri ile karşılaştırılmıştır. Lojistik regresyon modelleri kullanılarak kırılğanlık tahminlerinde bu parametrelerin duyarlılığı ve özgüllüğü değerlendirilmiştir. Lojistik regresyon modelleri, kırılğanlığı %100 duyarlılık ve özgüllük ile, ön-kırılğanlığı ise %87 duyarlılık ve %95 özgüllük ile tahmin edebilmiştir. Bu sonuçlar, önerilen üst ekstremite hareket değerlendirme yönteminin, kırılğanlık tespiti için son derece hassas ve spesifik olduğunu göstermektedir.

Parvaneh ve arkadaşları [22], toplulukta yaşayan yaşlı bireylerde kırılğanlık durumunu belirleyebilmek için giyilebilir teknoloji kullanımını araştırmıştır. Günlük yaşam aktiviteleri sırasında gerçekleştirilen hareket geçişlerinin, kırılğanlık durumunu izlemek için kullanılabilirliğini değerlendirmiştir. Yaş ortalaması  $78 \pm 8$  olan 120 yaşlı bireyin günlük fiziksel aktiviteleri, göğüs bölgesine takılan bir giyilebilir sensör ile 24 saat boyunca izlenmiş ve elde edilen veriler istatistiksel analiz metotlarıyla analiz edilmiştir. Ön-kırılğan ve kırılğan katılımcılarda hareket geçiş sayısının kırılğan olmayan bireylere nazaran daha az olduğu görülmüştür.

Razjouyan ve arkadaşları [23], toplulukta yaşayan yaşlı ve kırılğan bireylerde kırılğanlık durumunu tespit etmek için giyilebilir sensörlerin kullanımını incelemiştir. 60 yaş ve üzeri 153 yaşlı bireye, Fiziksel aktivite ve davranışları izlemek için 48 saat boyunca göğüs bölgesine takılan bir kolye sensör verilmiştir. Sensörden elde edilen verilerin analizi, lojistik regresyon modeli ve ROC eğrisi kullanılarak gerçekleştirilmiştir. Sonuçlar giyilebilir bir sensör kullanarak günlük ev içi izleme yoluyla ön-kırılğanlığın tanımlanabileceğini göstermiştir.

Straiton ve arkadaşları [24], tüketici düzeyinde aktiviteleri izlemek için kullanılan cihazların, yaşlılarda adım sayısı ve aktivite süresini ne kadar doğru ölçebildiğini ve bu cihazların araştırma amaçlı veya kişisel kullanım için uygun olup olmadığını değerlendirmek için geçerliliğini ve güvenilirliğini incelemiştir. 2014-2018 yılları arasında yayınlanmış yedi gözlemsel çalışmayı incelemiştir. Çalışmalarda, 65 yaş üstü 290 katılımcının fiziksel aktiviteleri, tüketici düzeyinde giyilebilir cihazlar kullanılarak izlenmiştir. Geçerlilik, referans cihazlar veya altın standart doğrulama yöntemleri ile karşılaştırıldığında korelasyon katsayıları ve yüzde hata ile değerlendirilmiştir. Güvenilirlik ise, adım sayısı ve aktivite süresi arasındaki farklar göz önünde bulundurularak ölçülmüştür. Tüketici düzeyindeki bu cihazlar, toplulukta yaşayan yaşlı yetişkinlerde fiziksel aktiviteyi izlemek için geçerli ve güvenilir bir seçenek olarak değerlendirmiştir.

Lee ve arkadaşları [25], hastane ortamında yatağa bağlı geriatri hastalarında kırılğanlık durumunu tespit etmek için tek bir bileğe takılan sensör kullanarak geliştirilen otomatik bir algoritmanın doğruluğunu incelemiştir. Yaş ortalaması 78.9 olan 100 geriatri hastası Kırılğanlık Metre adı verilen bir cihazla, bileğe takılan tek bir sensör kullanarak 20 saniyelik hızlı dirsek fleksiyon-ekstansiyon testi gerçekleştirmiş ve verilerin analizi için geliştirilen algoritmalar kullanılmıştır. Sonuçlar, tek bir bileğe takılan sensörle yapılan ölçümlerin, kırılğanlık durumunu belirlemede yeterince hassas olduğunu ve özellikle hastane ortamında kullanılabilirliğini göstermektedir.

Kırılğanlık üzerine yapılan çalışmaların belki de en önemli alanlarından biri bu tezin de konusu olan hızlı tespiti için metotlar geliştirmektir. Bu alanda hastalara giydirilen sensör ve veri işleme metotlarıyla hızlı ve doğru bir tahminlenme yapılması amaçlanmaktadır.

Toosizadeh ve arkadaşları [26], hastaneye yatırılmış yaşlı travma hastalarında kırılabilirlik durumunu belirlemek için yenilikçi bir üst ekstremite hareket değerlendirme yönteminin (UEF) doğruluğunu ve uygulanabilirliğini incelemiştir. 65 yaş ve üzeri, travmatik yer seviyesinde düşme nedeniyle hastaneye yatırılan 101 yaşlı bireye 20 saniyelik hızlı dirsek fleksiyonu testi sırasında UEF sistemini kullanarak ölçüm yapılmıştır. Ölçülen verilerin analizleri için çoklu regresyon modelleri kullanılmıştır. Kırılabilir bireylerde dirsek fleksiyon hızı, güç ve esneklik gibi parametreler, kırılabilir olmayan bireylere kıyasla anlamlı derecede daha düşük bulunmuştur. Ayrıca, UEF parametreleri ile son bir yıldaki düşme sayısı arasında anlamlı korelasyonlar tespit edilmiştir.

Hassler ve arkadaşları [27], kırılabilirlik sendromunu tahmin etmek ve risk faktörlerini keşfetmek için tıbbi veri ön işleme yöntemleri ve çeşitli tahmin modelleri geliştirmeye çalışmışlardır. 474 katılımcıdan elde edilen 284 sağlık parametresini içeren Toledo Study for Healthy Aging (TSHA) verileri kullanılmıştır. Katılımcılar, Fried kırılabilirlik kriterlerine göre kırılabilir ve kırılabilir olmayan olarak sınıflandırılmıştır. Çalışmada, eksik verilerin tamamlanması, veri ön işleme ve risk faktörlerinin belirlenmesi için çeşitli veri madenciliği ve makine öğrenme teknikleri kullanılmıştır. İkili sınıflandırma modelleri (SVM, LDA, CART gibi) geliştirilmiş ve bu modellerin performansı değerlendirilmiştir. Kullanılan modeller arasında Support Vector Machine (SVM) en yüksek doğruluğa (%78,31) sahip olmuştur. Çalışma, eksik ve dengesiz tıbbi verilerin kullanılarak kırılabilirlik sendromunu tahmin edebilecek modellerin geliştirilmesinin mümkün olduğunu göstermiştir.

Le Pogam ve arkadaşları [28], Fried kırılabilirlik fenotipini öngörmek ve yaşlı hastalarda bir yıl boyunca hastane taburculuk verileri kullanarak elektronik bir kırılabilirlik skoru (e-FS) geliştirmek ve e-FS'nin Fried kırılabilirlik fenotipini ve ilgili olumsuz sağlık sonuçlarını tahmin etme yeteneğini incelemişlerdir. Geliştirme ve iç doğrulama aşamasında, Lausanne Üniversitesi Hastanesi'nde 2004-2015 yılları arasında hastanede yatmış 469 katılımcının verilerini kullanmıştır. e-FS modeli, yaş, cinsiyet, son taburcu olma süresi ve e-FS skorunu içeren dört değişken ile geliştirilmiştir. Modelin performansı, lojistik regresyon modelleri ve diğer tahmin modelleri kullanılarak değerlendirilmiştir. e-FS modeli, Fried kırılabilirlik fenotipini öngörmeye ve sağlık sonuçlarını tahmin etmede başarılı olmuştur.

Leme ve arkadaşları [29], toplulukta yaşayan orta yaşlı ve yaşlı yetişkinlerde gelecekteki kırılabilirlik durumunu tahmin etmek için ML modellerinin etkinliğini ve özellikle dengeli ve dengesiz veri setleri kullanılarak eğitilen ML modellerinin performansını

karşılaştırmayı ve kırılabilirliği tahmin etmede en önemli değişkenleri belirlemeyi hedeflemektedir. Kırılabilir olmayan 4378 katılımcı değerlendirilmiş ve 4 yıl sonra yeniden kırılabilirliği fenotipi açısından değerlendirilmiştir. Sosyal, klinik ve psikososyal değişkenler kullanılarak, lojistik regresyon, Random Forest (RF), SVM, Sinir Ağı (NNET), K-en yakın komşu (K-NN), ve Naive Bayes sınıflandırıcısı gibi ML modelleri eğitilmiştir. Dengesiz veri setlerindeki dengesizlikleri düzeltmek için SMOTE yöntemiyle veri setleri dengelenmiştir. Dört yıllık takip sonunda, başlangıçta kırılabilir olmayan katılımcıların %7,9'u kırılabilir hale gelmiştir. Dengelenmiş veri setiyle eğitilen modellerde en iyi performansı RF modeli göstermiştir. Yaş, sandalye kalkma testi, hane halkı zenginliği, denge problemleri ve öz-değerlendirilmiş sağlık durumu, kırılabilirliği tahmin etmede en önemli değişkenler olarak belirlenmiştir.

Szczepanowski ve arkadaşları [30], kırılabilirliği tahmin etmek için kullanılan çeşitli ML algoritmalarını değerlendirerek, bu algoritmaların kalp yetmezliği hastalarında kırılabilirliği durumunu öngörmeye ne kadar etkili olduğunu ortaya koymayı hedeflemektedir. Kullanılan algoritmalar arasında SVM, Decision Tree, RF ve AdaBoost bulunmaktadır. Bu algoritmalar, Tilburg Frailty Indicator (TFI) anketine dayanan veriler kullanılarak eğitilmiş ve test edilmiştir. SVM, Decision Tree, RF ve AdaBoost algoritmalarının kırılabilirliği tahmininde yüksek doğruluk oranları sağladığını göstermiştir. AdaBoost algoritması, %100'e yakın bir doğruluk oranıyla en iyi performansı göstermiştir. Ayrıca, ML modelleri kullanılarak elde edilen sonuçlar, Fried kırılabilirlik fenotipinin fiziksel, psikolojik ve sosyal bileşenlerinin önemini ortaya koymuş ve bu bileşenlerin kırılabilirliği durumunu tahmin etmede önemli rol oynadığını göstermiştir.

Mizuguchi ve arkadaşları [31], yaşlı kalp yetmezliği hastalarında klinik kırılabilirlik ölçeğini otomatik olarak değerlendirmek için ML ve derin öğrenme tekniklerinin uygulanabilirliğini araştırmışlardır. 75 yaş ve üzeri semptomatik kronik kalp yetmezliği tanısı alan 417 yaşlı hasta değerlendirilmiştir. Yürüyüş analizi, akıllı telefon kamerası kullanılarak derin öğrenme tabanlı bir tahmin kütüphanesi aracılığıyla yapılmıştır. Yürüyüş parametreleri ve 128 anahtar özellik, Light Gradient Boosting Machine (LightGBM) modeli kullanılarak tahmin edilmiştir. LightGBM modeli gerçek ve tahmin edilen klinik kırılabilirlik ölçeği arasında mükemmel bir uyum göstermiştir.

Chen ve arkadaşları [32], KOAH hastalarında kırılabilirliği daha etkili bir şekilde tespit edebilmek için risk faktörlerini belirleyip bu faktörler üzerine kurulu bir tahmin modeli

oluşturmaya çalışmışlardır. 2018 Çin Sağlık ve Emeklilik Boylamsal Çalışması (CHARLS) veri tabanından elde edilen veriler analiz edilmiştir. Depresyon, sigara içme durumu, cinsiyet, sosyal aktiviteler, dislipidemi, astım ve yaşanılan yer gibi toplamda 34 değişkeni incelemiştir. Değişken seçimi için ridge regressor, XGBoost ve RF regressor yöntemleri uygulanmıştır. Logistic regression, SVM, multilayer perceptron (MLP), LightGBM, XGBoost, RF ve KNN gibi ML tekniği kullanılarak en iyi model belirlenmiştir. XGBoost modeli en yüksek doğruluğa sahip model olarak belirlenmiştir. Test setinde XGBoost modeli, %91.5 doğruluk, %87.3 duyarlılık ve %91.1 özgüllük oranlarına ulaşmıştır. Depresyon, sigara içme, cinsiyet, sosyal aktiviteler, dislipidemi, astım ve yerleşim yeri, kırılabilirliği gelişiminde en önemli öngörücü faktörler olarak belirlenmiştir. Bu bulgular, geliştirilen ML tabanlı tahmin modelinin KOAH hastalarında kırılabilirliği riskini değerlendirmek için kullanışlı bir araç olabileceğini göstermektedir.

Erdaş ve Ölçer [33], yaşlanan nüfus için önemli bir sorun haline gelen kırılabilirliğin derecelendirilmesi için ML modellerinin kullanımını ve etkinliğini araştırmışlardır. KNN, Karar Ağacı, Rastgele Orman, Ekstra Ağaçlar ve CatBoost modelleri kullanılmıştır. Bu modeller, TUG (Timed Up and Go) testi sırasında toplanan ivmeölçer sinyallerinden elde edilen zaman alanı öznelikleri ile beslenmiştir. Veriler, açık kalp cerrahisi sonrası kardiyak rehabilitasyon programına katılan kırılabilir hastalardan elde edilmiştir. Deneysel sonuçlar, 10 kat çapraz doğrulama yaklaşımı kullanılarak değerlendirilmiş ve regresyon performansı çeşitli metriklerle ölçülmüştür. Ekstra Ağaçlar modeli, en iyi performansı göstererek kırılabilirlik derecesi tahmini için en başarılı model olarak belirlenmiştir. Rastgele Orman yöntemi de benzer şekilde başarılı sonuçlar vermiştir. KNN ve CatBoost modelleri ortalama performans sergilerken, Karar Ağacı modeli tatmin edici olmayan sonuçlar vermiştir.

### 3. YÖNTEM VE MATERYAL

Bu başlık altında öncelikle veri seti ardından ise kullanılan yöntem ve metotlar tanıtılmaktadır. Tüm metot ve yöntemlerin tanıtılmasının ardından tüm çalışma detaylıca açıklanmaktadır.

#### 3.1. Veri Seti

Çalışmada kullanılan veriler, PhysioNet tarafından sağlanan halka açık veri kümesinden elde edilmiştir [34]. 19 Kasım 2020 ile 3 Ocak 2022 tarihleri arasında Kulautuva Rehabilitasyon Hastanesi Kaunas Klinikleri'ne (Kulautuva, Litvanya) gelen açık kalp ameliyatı geçirmiş hastalar üzerinden veriler toplanmıştır. Yatılı kalan hastalar için rutin bir işlem olarak sağlık personelleri, kırılganlık derecesini, Edmonton kırılganlık ölçeğine (EFS) göre değerlendirmiştir. Değerlendirmeler sonucunda 80 hastalık bir veri seti elde edilmiştir.

EKG sinyalleri 130 Hz'de ve ivmeölçer sinyalleri 200 Hz'de örneklenmiş olup, sinyaller Polar H10 giyilebilir cihazı (Polar Electro OY, Kempele, Finlandiya) kullanılarak elde edilmiştir. Cihaz, göğüs altına yerleştirilmiş olup, sinyaller gerçek zamanlı olarak Bluetooth bağlantısı aracılığıyla bir akıllı telefona aktarılmıştır. İzleme süresi boyunca, hastalar akıllı telefonu üst kollarına sarılı bir tutucuda taşımış ve egzersiz testinin başlangıcını bir mobil uygulama ile kaydetmiştir.

Bu çalışmada, hastalar üzerinde bazı fiziksel testler gerçekleştirmiştir. Bunlar:

- **Veloergometri:** Hastanın artan yükte bir sabit bisiklet üzerinde pedal çevirdiği bir testtir.
- **6 Dakika Yürüme Testi (6MWT):** Kişinin düz bir yüzeyde 6 dakika boyunca yürüdüğü ve kat edebildiği mesafenin ölçüldüğü bir testtir.
- **Merdiven Çıkma (Stair):** 12 basamaklı bir merdiveni yardımsız çıkan hastanın klinik verilerinin ölçüldüğü bir testtir.
- **Zamanlı Kalk ve Git Testi (TUG):** Bir kişinin sandalyeden kalkıp 3 metre yürüyüp geri döndüğü ve oturduğu, sürenin kaydedildiği bir testtir.
- **Gaint Analysis:** Bir koşu bandı üzerinde yürüyüş ve duruş analizinin yapıldığı bir testtir. Bu test esnasında adım frekansı, uzunluğu, süresi gibi değerler ölçülür.

Her fiziksel testten önce ve sonra, hastalar en az 3 dakika oturma pozisyonunda dinlenmişlerdir.

### 3.2. Veri Ön İşleme

Çalışmada, 80 hastalık veri setinde, 1 hasta kaydındaki adreslemelerin hatalı olduğu düşünüldüğünden kayıt veri setinden çıkarılarak 79 hasta kaydı üzerinden çalışmaya devam edilmiştir. Zaman serilerinden elde edilen verilerin kullanılabilmesi için anlamlı veriler haline getirilmesi bu verilerden öz nitelikler çıkarılması gerekmektedir. Çıkarılan bu öz niteliklerin kullanılmasıyla verinin anlamlı hale getirilmesi ve ilgili ML'nin daha doğru bir sonuç vermesi amaçlanmaktadır [35]. Teze konu bu çalışmada aşağıda bahsi geçen 14 metot kullanılmıştır:

- **Ortalama (Mean):** Ortalamayı, bir veri setindeki tüm değerlerin toplamının, veri sayısına bölünmesi olarak tanımlayabiliriz. Ortalamalar, veri setinin merkezini veya genel eğilimini gösterir.
- **Standart Sapma (Standard Deviation):** Standart sapma, veri setindeki verilerin ortalamadan ne kadar uzaklaştığını gösterir. Verilerin ne kadar yaygın olduğunu ölçmek için kullanılır.
- **Medyan (Median):** Medyan, bir veri setini sıraladıktan sonra tam ortada kalan değerdir. Veri setinde aşırı büyük veya küçük değerlerin etkisini ortadan kaldırır.
- **Çeyrekler Arası Aralık (Interquartile Range - IQR):** Çeyrekler arası aralık, veri setinin %25'lik dilimi ile %75'lik dilimi arasındaki farkı temsil eder. Aşırı uç değerlerden etkilenmeden verilerin yayılımını gösterir.
- **Minimum Değer (Minimum Value):** Minimum değer, veri setindeki en küçük değeri gösterir.
- **Maksimum Değer (Maximum Value):** Maksimum değer, veri setindeki en büyük değeri gösterir.
- **Maksimum ve Minimum Değer Arasındaki Fark (Range):** Maksimum değer ile minimum değer arasındaki farktır. Bu fark, veri setinin ne kadar geniş bir aralıkta dağıldığını gösterir.
- **Ortalama Mutlak Sapma (Mean Absolute Deviation - MAD):** Ortalama mutlak sapma, her bir veri noktasının ortalamadan ne kadar saptığını gösterir. Tüm sapmaların mutlak değerlerinin ortalaması alınır.
- **Ortanca Mutlak Sapma (Median Absolute Deviation - MedAD):** Ortanca mutlak sapma, her bir veri noktasının medyandan sapmalarının medyanını temsil eder. Medyanın etrafındaki yayılmayı gösterir.

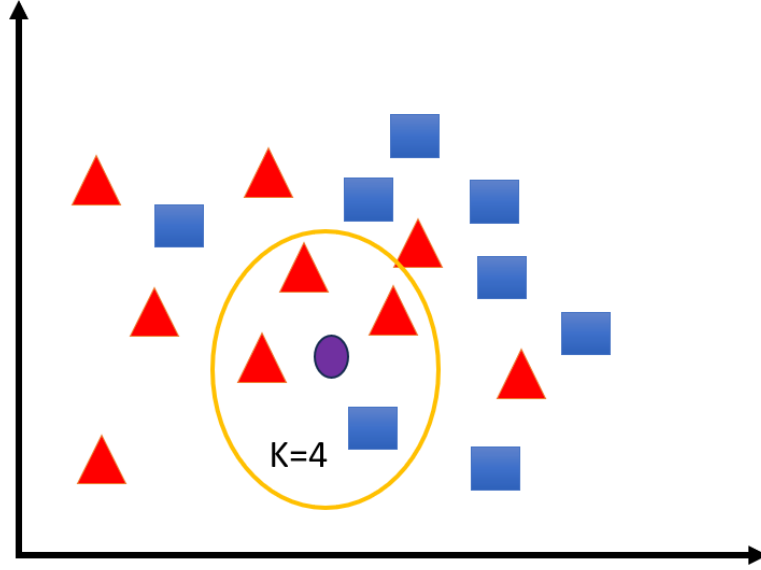
- **Ortalamanın Üstündeki Değer Sayısı (Count Above Mean):** Bir veri setinde ortalamadan daha büyük olan veri noktalarının sayısını ifade eder.
- **Ortalamanın Altındaki Değer Sayısı (Count Below Mean):** Bir veri setinde ortalamadan daha küçük olan veri noktalarının sayısını ifade eder.
- **Lokal Tepe Sayısı (Number of Local Peaks):** Lokal tepe sayısı, veri setinde komşu verilerden daha büyük olan noktaların (lokal maksimumların) sayısını ifade eder. Bu, verinin dalgalanma seviyesini gösterir.
- **Çarpıklık (Skewness):** Çarpıklık, bir veri setinin asimetrisini ölçer. Eğer dağılım sağa doğru uzun bir kuyruğa sahipse pozitif, sola doğru ise negatif çarpıklık söz konusu olur.
- **Basıklık (Kurtosis):** Basıklık, bir veri setinin dağılımının sivriliğini veya düzlüğünü ölçer. Yüksek basıklık, verilerin tepeye yakın yoğunlaştığını, düşük basıklık ise daha yayvan bir dağılım olduğunu gösterir.

Bu özellikler, veri setlerinin genel yapısını ve davranışını analiz etmek için kullanılan temel istatistiksel ölçütlerdir. Verilerin dağılımı, merkezi eğilimleri, yayılımı ve simetrisini anlamak için bu ölçütler sıklıkla kullanılır.

### 3.3. Kullanılan Algoritmalar

#### 3.3.1. K-En Yakın Komşu (KNN) Algoritması

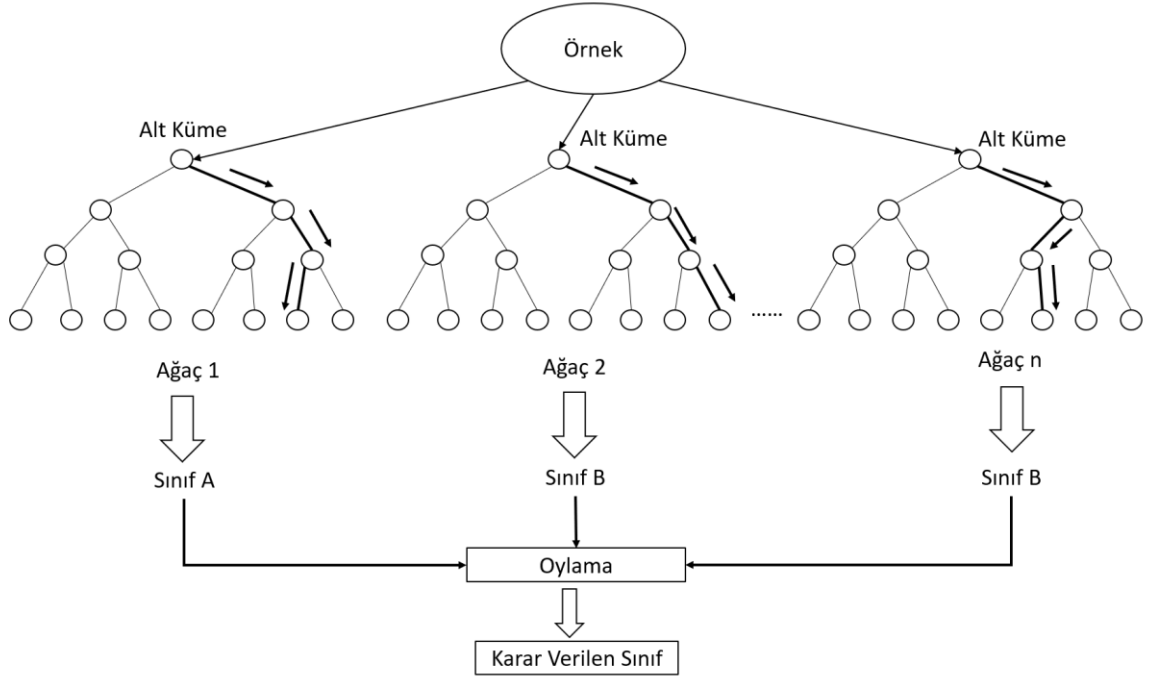
KNN yani K-En Yakın Komşu (K-Nearest Neighbors) algoritması ML metotları içerisinde özellikle sınıflandırma problemleri içinde oldukça popüler ve çokça tercih edilen bir algoritmadır. KNN her bir veri noktasını, belirli bir mesafe metriğine göre “k” adet en yakın komşusuyla ilişkilendirir. Bu mesafe Euclidean, Chebyshev, Manhattan gibi mesafeler ile ölçülür. Algoritmanın temel parametresi “k” parametresidir ve bu, komşu olarak seçilecek veri noktalarının sayısını ifade eder [36]. “k” küçük seçilirse, tahmin sadece en yakın birkaç veri noktasına dayanır, bu da genellikle veri noktalarının yakın çevresi için daha doğru sonuçlar verir. Ancak k=1 ise aşırı öğrenme (overfitting) olasılığı oldukça yüksektir. Aşağıdaki Şekil 3.1.’de görülebileceği gibi öznitelik uzayındaki bir nokta k=4 seçilmiş ve en yakın komşularından 4’ü seçilmiş ve kırmızı üçgen olarak sınıflandırılmıştır.



Şekil 3.1. K=4 için örnek KNN

### 3.3.2. Rastgele Orman (RF) Algoritması

Rastgele Orman (RF-Random Forests) algoritması, Leo Breiman tarafından 2001 yılında önerilen ve ML’de kullanılan güçlü bir sınıflandırma ve regresyon yöntemidir. RF, birden fazla karar ağacının birleştirilmesiyle oluşturulan bir modeldir. Her bir ağaç, eğitim verilerinin rastgele bir alt kümesi üzerinde eğitilir ve her düğümde, rastgele seçilmiş bir özellik alt kümesi kullanılarak bölünme işlemi gerçekleştirilir [37]. Bu, ağaçların bağımsız ve çeşitli olmasını sağlar. Birden çok karar ağacının bir araya gelmesi nedeniyle aşırı öğrenme sorunu yaşamaması bu algoritmanın en güçlü yönlerinden biridir. Şekil 3.2.’de görüldüğü gibi her ağaç, eğitim setinden rastgele örneklenmiş verilerle ve rastgele seçilmiş özelliklerle oluşturulur. Bu sayede ağaçlar arasındaki korelasyonu azaltarak, ormanın genelleme yeteneğini artırır. Sınıflandırma problemlerinde her bir ağaç kendi sonucunu verir ve en çok ağacın onayladığı sonuç seçilir. Regresyon görevlerinde ise, her bir ağacın tahminlerinin ortalaması alınarak son tahmin elde edilir [37].



Şekil 3.2. Rastgele Orman Sınıflandırma yapısı [38]

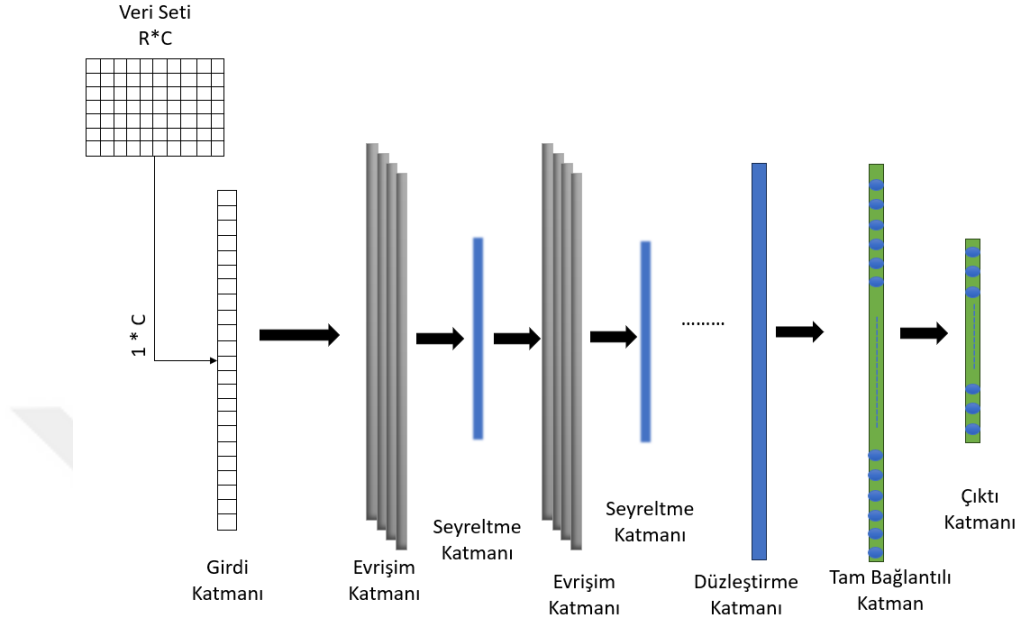
### 3.3.3. Ekstrem Gradyan Artırma (XGBoost) Algoritması

Ekstrem Gradyan Artırma (XGBoost-Extreme Gradient Boosting), denetimli öğrenme (supervised learning) problemleri için geliştirilmiş, güçlü ve esnek bir ML algoritmasıdır. XGBoost, temel olarak, gradyan artırma (gradient boosting) yönteminin optimize edilmiş ve genişletilmiş bir versiyonudur. Gradyan artırma, karar ağaçları gibi bir dizi modelin, hataları azaltarak birbirini takip edecek şekilde eğitildiği bir topluluk öğrenme tekniğidir. Her bir yeni modelin, mevcut modelin tahmin hatalarını öğrenmeye çalıştığı ve her eklenen karar ağacıyla birlikte daha doğru sonuç veren bir algoritmadır. Aynı zamanda her yeni ağaç eklenirken modelin toplam hatasını azaltmak için her bir ağaçta mevcut olan ağırlıkları güncellemektedir. Aşırı öğrenmenin önüne geçmek için karmaşıklığı kontrol altında tutar. Paralel veri işlemleri ve eksik verileri otomatik olarak işleme kabiliyeti ile oldukça avantajlı bir algoritmadır [39].

### 3.3.4. Tek Boyutlu Evrişimli Sinir Ağı (1D-CNN) Algoritması

Tek Boyutlu Evrişimli Sinir Ağı (1D-CNN-1D Convolutional Neural Network), tek boyutlu veriler üzerinde çalışan bir tür yapay sinir ağıdır. Genellikle zaman serisi verileri, ses verileri, sinyaller veya metinler gibi tek boyutlu veriler üzerinde kullanılır. Bu tür ağlar, verilerdeki özellikleri ve desenleri çıkarmak için bir dizi filtre veya çekirdek kullanarak

evrişim (convolution) işlemleri gerçekleştirir. Adaptif öğrenme, hız, verimlilik, öznelik çıkarımı yapabilme gibi hususlardaki avantajları sayesinde oldukça tercih edilen bir algoritmadır [40]. Şekil 3.3. Örnek 1D-CNN Modeli [41] verilmiştir.



Şekil 3.3 Örnek 1D-CNN Modeli [41]

### 3.3.5. Naive Bayes Algoritması

Olasılık temelli ve özellikler arasında bağımsızlık varsayımına dayanan basit bir ML algoritmasıdır. Bayes teoremini kullanarak çalışır ve genellikle sınıflandırma görevlerinde kullanılır. Naive Bayes sınıflandırıcısının temel avantajı, özellikle büyük özellik alanlarında etkili olması ve hesaplama açısından verimli olmasıdır [42]. Eşitlik 3.1’de verilen  $P(\text{Sınıf}|\text{Özellikler})$  özellikler göz önünde bulundurulduğunda sınıfın olasılığı,  $P(\text{Sınıf})$  sınıfın olasılığı,  $P(\text{Sınıf}|\text{Özellikler})$  sınıf verildiğinde özelliklerin olasılığı ve  $P(\text{Özellikler})$  Özelliklerin toplam olasılığını ifade eder [42].

$$P(\text{Sınıf}|\text{Özellikler}) = \frac{P(\text{Sınıf}) * P(\text{Özellikler}|\text{Sınıf})}{P(\text{Özellikler})} \quad (3.1)$$

Hesaplama maliyetinin düşük olması nedeniyle hızlı ve görece daha az veri ile daha doğru sonuçlar sunabilme avantajlarına rağmen bağımsızlık varsayımından yani özelliklerin birbirinden bağımsız olduğu varsayımı gerçek dünyayla örtüşmemesi sebebiyle sonucu olumsuz yönde etkileyebilmektedir [43].

### 3.3.6. Çok Katmanlı Algılayıcı (MLP) Algoritması

Çok Katmanlı Algılayıcı (MLP-Multi-Layer Perceptron), yapay sinir ağlarının temel bir türüdür ve çoğunlukla sınıflandırma ve regresyon gibi problemler için kullanılır. Bu model, giriş katmanı, bir veya daha fazla gizli katman ve çıkış katmanı olmak üzere birkaç katmandan oluşur ve her katmandaki nöronlar, bir sonraki katmandaki tüm nöronlarla bağlantılıdır. Giriş katmanındaki nöronlar, modelin işlediği veri özelliklerini temsil ederken, gizli katmanlar veriyi işleyerek öğrenme sürecine katkıda bulunur. Çıkış katmanı, modelin elde ettiği tahmin sonuçlarını oluşturur. MLP’de nöronlar, doğrusal olmayan aktivasyon fonksiyonları kullanarak veriyi işler. Modelin eğitim süreci, geri yayılım algoritmasıyla gerçekleştirilir; bu algoritma, modelin tahmin hatalarını minimize etmek için ağırlıkları günceller [44].

### 3.4. Performans Değerlendirmesi İçin Kullanılan Metrikler

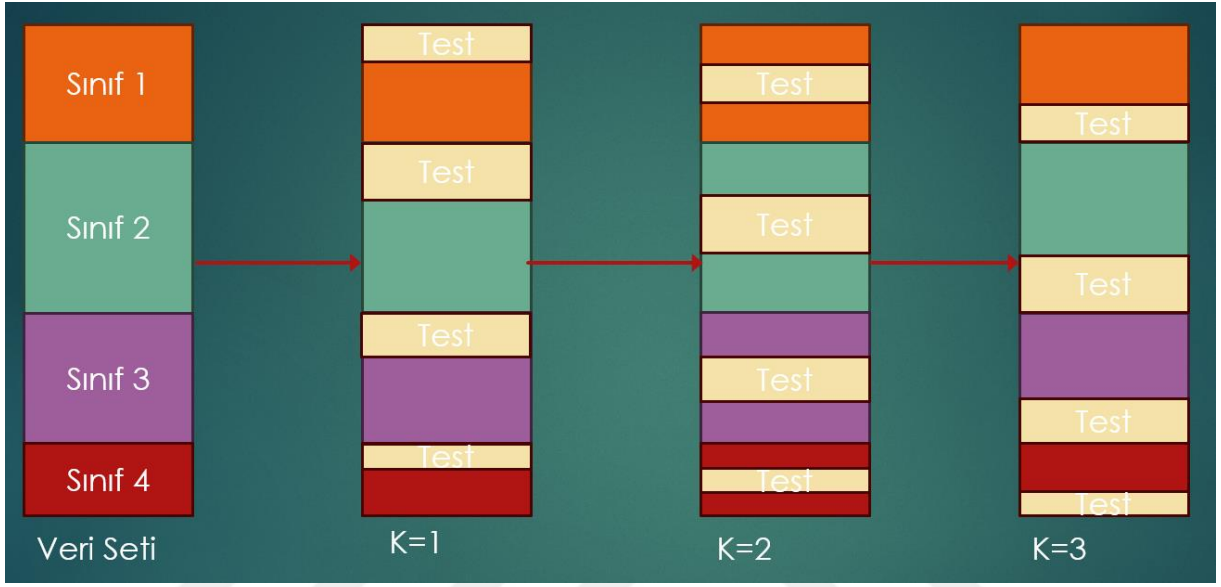
Bu tez çalışmasında sınıflandırma problemlerinde kullanılan Doğruluk (accuracy), kesinlik (precision), duyarlılık (Recall), F1 skoru (F1 score) ve Matthews Korelasyon Katsayısı (MCC-Matthews Correlation Coefficient) değerlendirme metrikleri kullanılmaktadır. Aynı zamanda eğitim ve test kümelerinin bölünme yöntemlerinden Stratified K-fold yöntemi kullanılmaktadır.

#### 3.4.1. Stratifli K-Fold

Katmanlı K-kat çapraz doğrulama (Stratified K-Fold Cross-Validation) (SKCV), özellikle sınıflandırma problemlerini çözmek için kullanılan bir çapraz doğrulama yöntemidir. Bu yöntem, veri setindeki sınıf dağılımlarını her bir katlama (fold) içinde koruyarak, modelin daha dengeli ve tutarlı bir şekilde değerlendirilmesini sağlar. SKCV, veri setini K adet alt gruba ayırır ve her bir alt grup, sırasıyla doğrulama ve eğitim seti olarak kullanılır. Bu şekilde, modelin farklı veri bölümleri üzerinde nasıl performans gösterdiği detaylı bir şekilde analiz edilebilir [45].

SKCV'nin en önemli özelliği, basit çaprazlama yöntemini geliştirmesidir. K-Fold yönteminde veri seti rastgele bölünürken, SKCV'de bu bölünme işlemi sınıf dağılımları dikkate alınarak yapılır. Böylece, her bir katlamada, orijinal veri setindeki sınıf oranları korunur ve modelin, özellikle dengesiz sınıf dağılımlarına sahip veri setleri üzerinde daha tutarlı sonuçlar vermesi sağlanır [45].

Şekil 3.4.'te görülebileceği üzere, eşit dağılımlı olmayan 4 sınıftan oluşan bir veri seti verilmiş ve bu veri setinin eğitimi esnasında SKCV kullanılarak k=3 olarak seçilmiştir. Bu durumda her bir sınıf 3'e bölünecek ve bir test verisi yaratılacak bunun dışındaki veriler ise eğitimde kullanılacaktır. Bir başka deyişle, ana kümedeki sınıf dağılım oranları göz önüne alınarak her örneklemede ana kümedeki oranla eşit oranda veri bulunur.



Şekil 3.4. SKCV K=3

### 3.4.2. Doğruluk (Accuracy)

Doğruluk [46], Doğruluk (accuracy), ML modellerinin başarısını ölçmek için en sık kullanılan metriklerden biridir. Basitçe açıklamak gerekirse, doğru tahmin edilen sonuçların tüm tahminlere oranı olarak hesaplanır. Eşitlik 3.2'de doğruluğun matematiksel ifadesi gösterilmiştir.

$$Doğruluk = \frac{DP + DN}{DP + DN + YP + YN} \quad (3.1)$$

Doğruluk, karmaşıklık matrisi (confusion matrix) üzerinden türetilen bir metrik olup, modelin performansını değerlendirmek için oldukça önemlidir. Karmaşıklık matrisi, tahmin edilen ve gerçek sonuçlar arasındaki ilişkiyi görselleştirir ve dört temel olasılık üzerinden analiz yapılmasını sağlar:

- DP (Doğru Pozitif): Pozitif tahmin edilen sonucun gerçekten pozitif olması,
- YP (Yanlış Pozitif): Pozitif tahmin edilen sonucun aslında negatif olması,

- YN (Yanlış Negatif): Negatif tahmin edilen sonucun aslında pozitif olması,
- DN (Doğru Negatif): Negatif tahmin edilen sonucun gerçekten negatif olması anlamlarına gelir.

### 3.4.3. Kesinlik (Precision)

Kesinlik, doğruluk metriğinin yaratabileceği yanlışlığı tespit etmek için birlikte kullanılacak metriklerden biridir. Eşitlik 3.3'te görülebileceği üzere doğru tahmin edilen pozitif değerlerin gerçekteki pozitif değerlere olan oranını ölçmektedir.

Bu sayede dengeli olmayan bir veri setinde sadece pozitif değerler içinde bakarak daha doğru bir sonuç elde edilebilir [47].

$$Kesinlik = \frac{DP}{DP + YP} \quad (3.3)$$

### 3.4.4. Duyarlılık (Recall)

Duyarlılık, doğruluk metriğinin yaratabileceği illüzyonu tespit etmek için birlikte kullanılacak metriklerden bir diğeridir. Eşitlik 3.4'te görüleceği üzere doğru tahmin edilen pozitif değerlerin toplam pozitif tahminlere oranını ölçmektedir [47]. Kesinlik ile kullanıldığında doğruluğun yaratabileceği yanlışlığa engel olmadıkça önemli bir metriktir.

$$Duyarlılık = \frac{DP}{DP + YN} \quad (3.4)$$

### 3.4.5. F1 Skoru

Kesinlik ve Duyarlılık metriklerinin birlikte kullanılması modelin başarısının güvenilirliği açısından oldukça önemlidir. Ancak bu iki metriktir gelen sonuçların kıyaslanması, bir araya getirilerek bir anlam çıkarılması da başlı başına bir sorun teşkil eder. Bu nedenle bu iki metriğin harmonik bir ortalaması olan F1 Skor metriği kullanılır. Harmonik ortalamasının alınması oluşabilecek uç durumlarda bu metriğin yanıltıcı olmaması içindir. Örneğin aritmetik ortalama kullandığımızı ve duyarlılık değerinin 0

kesinlik deęerinin 1 ıktığı bir durumda F1 Skor 0.5 ıkararak yanıltıcı bir sonuç verecektir. F1 Skorunun matematiksel formülü Eşitlik 3.5'te verilmektedir [47].

$$F1\ Skor u = 2 * \frac{Kesinlik * Duyarlılık}{Kesinlik + Duyarlılık} \quad (3.5)$$

### 3.4.6. MCC (Matthews Korelasyon Katsayısı)

Dengeli olmayan bir veri setinde modelin başarısını ölçmek için en verimli metrik MCC'dir. Diğer metrikler tam anlamıyla güvenilir bir sonuç verememektedir [40]. Ancak MCC diğer metriklerin aksine karmaşıklık matrisindeki tüm deęerleri kullanarak modelin doğruluęunu test eden bir metrik olması sebebiyle tüm metrikler arasında en güvenilir olanıdır. MCC -1 ila +1 arası deęerler alır. MCC'nin 1 olması tamamen doğru tahminlerde bulunduęu, 0 olması rastgele tahminlerde bulunduęu, -1 olması ise tamamen yanlış tahminlerde bulunduęu anlamına gelir [48]. MCC'nin matematiksel formülü Eşitlik 3.6'da verilmektedir.

$$MCC = \frac{(DP * DN - YP * YN)}{\sqrt{(DP + YP) * (DP + YN) * (DN + YP) * (DN + YN)}} \quad (3.6)$$

### 3.5. Deneysel Kurulum

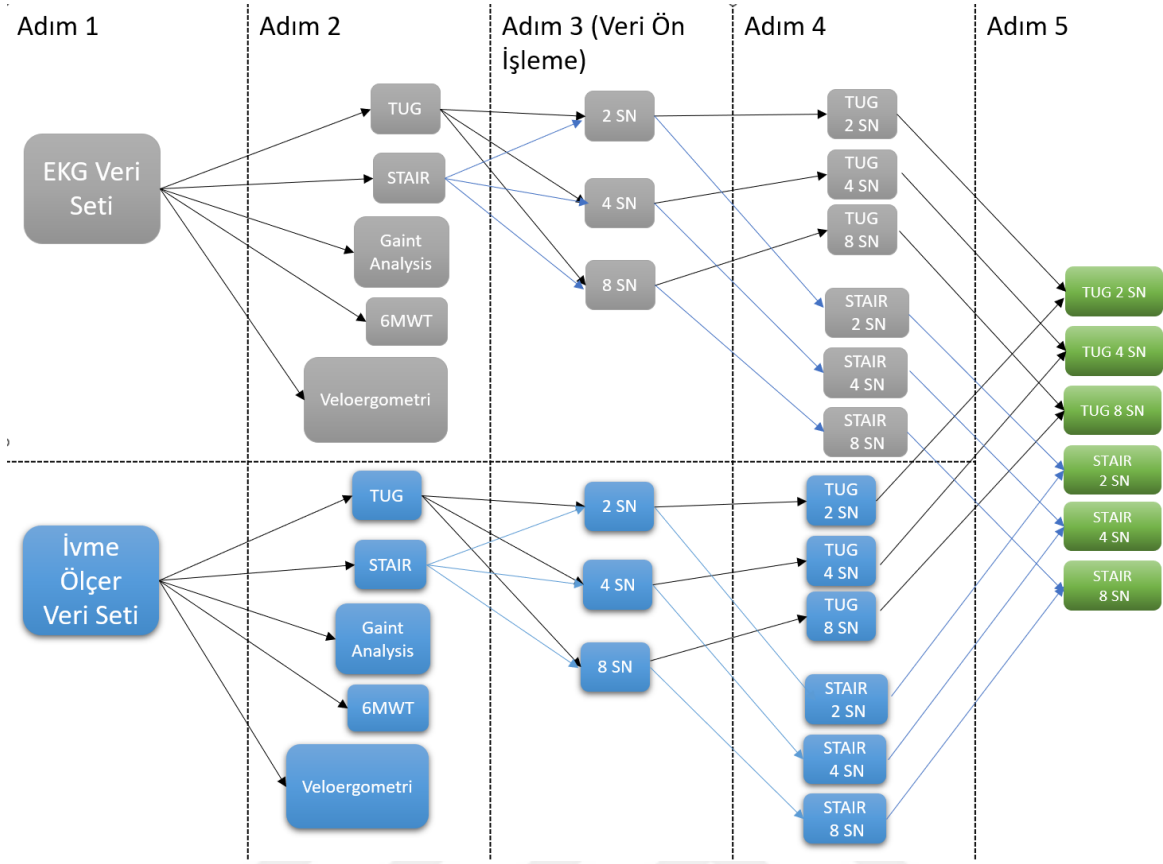
Yapılan alıřmada amaca yönelik olarak işlenmiş ve toplamda 45 ayrı veri seti ıkarılarak bu veri setleri üzerinde alıřılmıştır. Kırılğanlık öleęi olarak ise Edmonton Frail Scale (EFS) kullanılmıştır.

İlk olarak Tablo 3.1.'de verildięi gibi, her bir hastanın EFS deęeri alınarak 0 – 4 aralıęında saęlıklı, 5 – 6 aralıęında kırılğanlık başlangıcı, 7 – 8 aralıęında düşük düzeyde kırılğanlık, 9 – 10 aralıęında orta düzey kırılğanlık ve 11 üzeri (11 dahil) yüksek seviye kırılğanlığı ifade edecek 5 farklı sınıfa dahil edilmişlerdir.

**Tablo 3.1.** EFS skoruna göre Kırılgnlık sınıflandırmaları

| EFS Skoru | Sınıf                          |
|-----------|--------------------------------|
| 0-4       | Sağlıklı                       |
| 5-6       | Başlangıç Düzeyinde Kırılgnlık |
| 7-8       | Düşük Düzeyde Kırılgnlık       |
| 9-10      | Orta Düzey Kırılgnlık          |
| 11-15     | Yüksek Seviye Kırılgnlık       |

Sonrasında Şekil 3.5.'te gösterilen akış şemasına göre veri setleri oluşturulmuştur. Adım 1'de hastaların EKG ve İvmeölçer kayıtları okunmuştur. Bu adım sonucunda, Sensörlere göre hangi verinin hangi hastaya ait olduğu ve hastaların EFS sonuçlarını kaybetmemeye özen göstererek iki ayrı veri seti (EKG ve İvmeölçer) oluşturulmuştur. Adım 2'de EKG verileri için ayrı, İvmeölçer verileri için ayrı olacak şekilde testler özelinde veri setleri parçalanarak İvmeölçer verileri için 5, EKG verileri için 5 farklı veri seti elde edilmiştir. Bu işlemler yapılırken yine verinin hangi hastaya ait olduğu bilgisi korunmuştur. Adım 3'te zamansal öznitelik çıkarımı yapılmıştır. Öznitelik çıkarımı yapılırken 2, 4 ve 8 saniyelik zaman ölçekleri kullanılmıştır. Zaman ölçekleri veri seti içerisindeki hasta bilgileri göz önünde bulundurularak çıkarılmış ve yine hangi verinin hangi hastaya ait olduğu bilgisi korunmuştur. Zaman ölçeği %50 zaman örtüşmesine izin verilecek şekilde uygulanmıştır. Adım 3 tamamlandığında Adım 4'te EKG için 15, İvmeölçer için 15 olmak üzere toplamda 30 farklı veri seti elde edilmiştir. Son olarak Adım 5'te EKG ve İvmeölçer verileri birleştirilerek bir bileşik veri seti yaratılmıştır. Bu işlem her bir test ve o teste ait zaman ölçeği aynı olacak şekilde yapılmış bu veriler birleştirilirken yine hasta bilgileri göz önünde bulundurulmuştur. Sonuç olarak, 15 İvmeölçer, 15 EKG ve 15 İvmeölçer ve EKG verilerinden oluşan toplamda 45 farklı veri seti elde edilmiştir.



Şekil 3.5. Veri Setleri Oluşturma Adımları

Oluşturulan 45 veri seti KNN, Naive Bayes, MLP, RF, XGBoost ve 1D-CNN olmak üzere altı farklı algoritmayı eğitmek ve test etmek için kullanılmıştır. Bu çalışmada SKCV için  $K=10$ ,  $shufle=True$  ve  $random\_Satate=101$  parametreleri seçilmiştir.

KNN algoritmasında en iyi “k” değeri ve uzaklık metriğinin bulunması için Grid Search kullanılmıştır. Bu yöntem ile verilen potansiyel k kümesindeki ve uzaklık metrikleri kümesindeki tüm elemanları dener ve verilen çapraz doğrulama metriğine göre doğrularak en iyi sonucu veren k ve uzaklık metriğini geri döndürür. Bu çalışmada potansiyel K kümesi 1,3,5,7,9,11,13,15 ve potansiyel uzaklık metrikleri Öklid, Manhattan, Chebyshev şeklindedir.

XGBoost ve Random Forest için karar ağacı sayısı 100 olarak seçilmiştir. MLP için ise hidden layer 100 maksimum iterasyon sayısı olarak 1000 ve  $random\_state$  101 olarak belirlenmiştir. Son olarak 1D-CNN için 512 filtre ve relu ile softmax aktivasyon fonksiyonları kullanılmıştır.

### 3.6. Deneysel Sonular

Her bir test zaman aralıęı ve kullanılan sensör (İvmeöler ve EKG) için ayrı ayrı uygulanan 6 algoritma için performans metrik sonuçlarından oluşan tablolar ařaęıda listelenmektedir.

#### a) Stair testi için alınan sonuçlar

Tablo 3.2. İvmeöler STAIR 2 saniye’de STAIR testi esnasında toplanan üç eksenli İvmeöler verisi ile 2 saniyelik zaman aralıęında oluşturulmuş veri seti için sonuçlar görölmektedir. En iyi sonucu %90 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduęu görölmektedir. Grid Search ile yapılan aramada KNN için “k” deęeri 13, mesafe ölçütü ise Manhattan olarak bulunmuřtur.

**Tablo 3.2.** İvmeöler STAIR 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doęruluk   | %79,60 | %22,90      | %92,00  | %93,40 | %74,80 | %79,70 |
| Kesinlik   | %79,80 | %35,50      | %92,10  | %93,50 | %75,60 | %80,10 |
| Duyarlılık | %79,60 | %22,90      | %92,00  | %93,40 | %74,80 | %77,00 |
| F1 Skoru   | %79,40 | %19,40      | %92,00  | %93,40 | %74,40 | %78,30 |
| MCC        | %70,50 | %8,60       | %88,60  | %90,50 | %63,80 | %70,80 |

Tablo 3.3. İvmeöler STAIR 4 saniye’de STAIR testi esnasında toplanan ivmeöler verisi ile 4 saniyelik zaman aralıęında oluşturulmuş veri seti için sonuçlar görölmektedir. En iyi sonucu %90 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduęu görölmektedir. Grid Search ile yapılan aramada KNN için “k” deęeri 7, mesafe ölçütü ise Manhattan olarak bulunmuřtur.

**Tablo 3.3.** İvmeöler STAIR 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doęruluk   | %82,50 | %22,50      | %92,60  | %93,40 | %72,80 | %80,10 |
| Kesinlik   | %82,60 | %36,40      | %92,60  | %93,50 | %73,90 | %80,40 |
| Duyarlılık | %82,50 | %22,50      | %92,60  | %93,40 | %72,80 | %77,20 |
| F1 Skoru   | %82,30 | %19,20      | %92,60  | %93,30 | %72,40 | %78,50 |
| MCC        | %74,70 | %7,90       | %89,40  | %90,50 | %60,80 | %71,30 |

Tablo 3.4. İvmeölçer STAIR 8 saniye’de STAIR testi esnasında toplanan ivmeölçer verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %89 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.4.** İvmeölçer STAIR 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %81,90 | %22,30      | %92,80  | %92,70 | %72,30 | %79,90 |
| Kesinlik   | %82,00 | %37,20      | %92,90  | %92,90 | %73,20 | %80,00 |
| Duyarlılık | %81,90 | %22,30      | %92,80  | %92,70 | %72,30 | %77,30 |
| F1 Skoru   | %81,80 | %18,70      | %92,80  | %92,70 | %71,60 | %78,40 |
| MCC        | %74,00 | %7,30       | %89,70  | %89,60 | %59,80 | %71,10 |

Tablo 3.5. EKG STAIR 2 saniye’de STAIR testi esnasında toplanan EKG verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %97 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.5.** EKG STAIR 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %86,30 | %24,10      | %98,30  | %98,40 | %84,80 | %89,30 |
| Kesinlik   | %86,40 | %57,80      | %98,30  | %98,40 | %85,00 | %88,80 |
| Duyarlılık | %86,30 | %24,10      | %98,30  | %98,40 | %84,80 | %88,60 |
| F1 Skoru   | %86,20 | %15,50      | %98,30  | %98,40 | %84,70 | %88,60 |
| MCC        | %80,40 | %15,80      | %97,60  | %97,70 | %78,30 | %84,70 |

Tablo 3.6. EKG STAIR 4 saniye’de STAIR testi esnasında toplanan EKG verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile RF ve XGBoost modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.6.** EKG STAIR 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %88,30 | %24,10      | %98,90  | %98,90 | %85,10 | %87,80 |
| Kesinlik   | %88,30 | %60,00      | %98,90  | %98,90 | %85,20 | %87,50 |
| Duyarlılık | %88,30 | %24,10      | %98,90  | %98,90 | %85,10 | %86,80 |
| F1 Skoru   | %88,20 | %16,10      | %98,90  | %98,90 | %85,00 | %87,00 |
| MCC        | %83,20 | %15,90      | %98,40  | %98,40 | %78,70 | %82,60 |

Tablo 3.7. EKG STAIR 8 saniye’de STAIR testi esnasında toplanan EKG verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile RF ve XGBoost modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 3, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.7.** EKG STAIR 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %89,00 | %25,50      | %99,00  | %98,90 | %86,10 | %86,70 |
| Kesinlik   | %89,10 | %56,20      | %99,00  | %98,90 | %86,20 | %87,40 |
| Duyarlılık | %89,00 | %25,50      | %99,00  | %98,90 | %86,10 | %84,60 |
| F1 Skoru   | %89,00 | %18,80      | %99,00  | %98,90 | %86,10 | %85,70 |
| MCC        | %84,30 | %17,30      | %98,50  | %98,40 | %80,10 | %81,00 |

Tablo 3.8. İvmeölçer & EKG STAIR 2 saniye’de STAIR testi esnasında toplanan ivmeölçer ile EKG verileri ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile RF ve XGBoost modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.8.** İvmeölçer & EKG STAIR 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %92,10 | %26,90      | %99,40  | %99,40 | %91,90 | %92,20 |
| Kesinlik   | %92,10 | %39,60      | %99,40  | %99,40 | %92,00 | %92,30 |
| Duyarlılık | %92,10 | %26,90      | %99,40  | %99,40 | %91,90 | %91,50 |
| F1 Skoru   | %92,00 | %21,00      | %99,40  | %99,40 | %91,90 | %91,80 |
| MCC        | %88,60 | %16,10      | %99,20  | %99,10 | %88,50 | %88,80 |

Tablo 3.9. İvmeölçer & EKG STAIR 4 saniye’de STAIR testi esnasında toplanan ivmeölçer ile EKG verileri ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile RF ve XGBoost modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.9.** İvmeölçer & EKG STAIR 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %93,20 | %26,90      | %99,50  | %99,50 | %91,30 | %92,90 |
| Kesinlik   | %93,30 | %41,50      | %99,50  | %99,50 | %91,40 | %93,00 |
| Duyarlılık | %93,20 | %26,90      | %99,50  | %99,50 | %91,30 | %92,30 |
| F1 Skoru   | %93,20 | %21,30      | %99,50  | %99,50 | %91,30 | %92,60 |
| MCC        | %81,70 | %16,30      | %99,30  | %99,20 | %77,40 | %81,30 |

Tablo 3.10. İvmeölçer & EKG STAIR 8 saniye’de STAIR testi esnasında toplanan ivmeölçer ve EKG verileri ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile RF ve XGBoost modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 3, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.10.** İvmeölçer & EKG STAIR 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %93,20 | %26,20      | %99,50  | %99,40 | %91,80 | %92,20 |
| Kesinlik   | %93,20 | %42,70      | %99,50  | %99,40 | %91,90 | %92,00 |
| Duyarlılık | %93,20 | %26,20      | %99,50  | %99,40 | %91,80 | %91,60 |
| F1 Skoru   | %93,10 | %20,50      | %99,50  | %99,40 | %91,80 | %91,80 |
| MCC        | %90,20 | %16,00      | %99,30  | %99,20 | %88,30 | %88,90 |

**b) TUG testi için alınan sonuçlar**

Tablo 3.11. İvmeölçer TUG 2 saniye’de TUG testi esnasında toplanan ivmeölçer verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %89 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 9, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.11.** İvmeölçer TUG 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %77,90 | %47,20      | %90,50  | %92,50 | %73,50 | %78,60 |
| Kesinlik   | %78,20 | %45,20      | %90,70  | %92,60 | %74,40 | %79,80 |
| Duyarlılık | %77,90 | %47,20      | %90,50  | %92,50 | %73,50 | %77,00 |
| F1 Skoru   | %77,80 | %42,20      | %90,50  | %92,50 | %73,40 | %78,10 |
| MCC        | %69,00 | %23,20      | %86,80  | %89,60 | %63,10 | %70,10 |

Tablo 3.12. İvmeölçer TUG 4 saniye’de TUG testi esnasında toplanan ivmeölçer verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %89 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.12.** İvmeölçer TUG 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %80,30 | %47,10      | %90,80  | %92,20 | %72,10 | %78,10 |
| Kesinlik   | %80,50 | %45,70      | %90,90  | %92,40 | %73,40 | %79,00 |
| Duyarlılık | %80,30 | %47,10      | %90,80  | %92,20 | %72,10 | %76,70 |
| F1 Skoru   | %80,20 | %42,20      | %90,80  | %92,20 | %71,70 | %77,60 |
| MCC        | %72,40 | %23,10      | %87,20  | %89,10 | %61,00 | %69,50 |

Tablo 3.13. İvmeölçer TUG 8 saniye ‘de TUG testi esnasında toplanan ivmeölçer verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %87 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.13.** İvmeölçer TUG 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %80,20 | %47,00      | %90,50  | %91,20 | %69,00 | %75,80 |
| Kesinlik   | %80,40 | %45,50      | %90,70  | %91,50 | %70,60 | %77,50 |
| Duyarlılık | %80,20 | %47,00      | %90,50  | %91,20 | %69,00 | %74,40 |
| F1 Skoru   | %80,10 | %42,20      | %90,50  | %91,20 | %68,50 | %75,50 |
| MCC        | %72,40 | %23,00      | %86,80  | %87,80 | %56,50 | %66,30 |

Tablo 3.14. EKG TUG 2 saniye ‘de TUG testi esnasında toplanan EKG ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.14.** EKG TUG 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %91,40 | %35,60      | %99,30  | %99,00 | %88,90 | %93,40 |
| Kesinlik   | %91,40 | %39,40      | %99,30  | %99,00 | %89,10 | %93,60 |
| Duyarlılık | %91,40 | %35,60      | %99,30  | %99,00 | %88,90 | %93,20 |
| F1 Skoru   | %91,40 | %30,30      | %99,30  | %99,00 | %88,90 | %93,40 |
| MCC        | %88,00 | %13,70      | %99,00  | %98,70 | %84,60 | %90,90 |

Tablo 3.15. EKG TUG 4 saniye ‘de TUG testi esnasında toplanan EKG ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %97 MCC oranı ile RF ve XGBoost modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.15.** EKG TUG 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %79,90 | %28,60      | %98,40  | %98,10 | %75,60 | %73,60 |
| Kesinlik   | %79,80 | %50,80      | %98,40  | %98,10 | %75,80 | %72,70 |
| Duyarlılık | %79,90 | %28,60      | %98,40  | %98,10 | %75,60 | %69,80 |
| F1 Skoru   | %79,60 | %17,70      | %98,40  | %98,10 | %75,20 | %70,60 |
| MCC        | %71,90 | %11,60      | %97,70  | %97,30 | %66,00 | %63,20 |

Tablo 3.16. EKG TUG 8 saniye ‘de TUG testi esnasında toplanan EKG ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.16.** EKG TUG 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %82,80 | %28,60      | %98,70  | %98,30 | %77,80 | %73,60 |
| Kesinlik   | %82,80 | %45,60      | %98,70  | %98,30 | %78,30 | %73,20 |
| Duyarlılık | %82,80 | %28,60      | %98,70  | %98,30 | %77,80 | %69,80 |
| F1 Skoru   | %82,70 | %17,70      | %98,70  | %98,30 | %77,50 | %70,60 |
| MCC        | %76,10 | %11,20      | %98,10  | %97,70 | %69,20 | %63,30 |

Tablo 3.17. İvmeölçer & EKG TUG 2 saniye ‘de TUG testi esnasında toplanan ivmeölçer ve EKG verileri ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.17.** İvmeölçer & EKG TUG 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %91,40 | %35,60      | %99,30  | %99,00 | %88,90 | %93,50 |
| Kesinlik   | %91,40 | %39,40      | %99,30  | %99,00 | %89,10 | %93,50 |
| Duyarlılık | %91,40 | %35,60      | %99,30  | %99,00 | %88,90 | %93,50 |
| F1 Skoru   | %91,40 | %30,30      | %99,30  | %99,00 | %88,90 | %93,50 |
| MCC        | %88,00 | %13,70      | %99,00  | %98,60 | %84,60 | %90,90 |

Tablo 3.18. İvmeölçer & EKG TUG 4 saniye ‘de TUG testi esnasında toplanan ivmeölçer ve EKG verileri ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.18.** İvmeölçer & EKG TUG 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %92,20 | %33,50      | %99,50  | %99,20 | %88,50 | %92,40 |
| Kesinlik   | %92,20 | %38,30      | %99,50  | %99,20 | %88,70 | %92,60 |
| Duyarlılık | %92,20 | %33,50      | %99,50  | %99,20 | %88,50 | %92,10 |
| F1 Skoru   | %92,20 | %28,00      | %99,50  | %99,20 | %88,50 | %92,30 |
| MCC        | %89,10 | %11,40      | %99,30  | %98,80 | %84,00 | %89,50 |

Tablo 3.19. İvmeölçer & EKG TUG 8 saniye’de TUG testi esnasında toplanan ivmeölçer ve EKG verileri ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 3, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.19.** İvmeölçer & EKG TUG 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %93,00 | %31,40      | %99,40  | %99,20 | %87,80 | %92,50 |
| Kesinlik   | %93,10 | %36,20      | %99,40  | %99,20 | %88,10 | %92,80 |
| Duyarlılık | %93,00 | %31,40      | %99,40  | %99,20 | %87,80 | %92,30 |
| F1 Skoru   | %93,00 | %25,20      | %99,40  | %99,20 | %87,80 | %92,50 |
| MCC        | %90,30 | %8,70       | %99,20  | %98,90 | %83,10 | %89,60 |

### c) Veloergometri testi için alınan sonuçlar

Tablo 3.20. İvmeölçer Veloergometri 2 saniye’de Veloergometri testi esnasında toplanan ivmeölçer verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %90 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.20.** İvmeölçer Veloergometri 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %84,10 | %23,30      | %91,20  | %93,40 | %73,60 | %83,20 |
| Kesinlik   | %84,20 | %36,70      | %91,20  | %93,50 | %74,10 | %82,50 |
| Duyarlılık | %84,10 | %23,30      | %91,20  | %93,40 | %73,60 | %81,20 |
| F1 Skoru   | %84,00 | %18,70      | %91,20  | %93,30 | %73,10 | %81,70 |
| MCC        | %76,40 | %10,30      | %87,00  | %90,20 | %60,50 | %75,30 |

Tablo 3.21. İvmeölçer Veloergometri 4 saniye’de Veloergometri testi esnasında toplanan ivmeölçer verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %89 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 3, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.21.** İvmeölçer Veloergometri 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %86,70 | %21,20      | %91,90  | %93,10 | %72,30 | %81,50 |
| Kesinlik   | %86,80 | %35,60      | %92,00  | %93,20 | %73,00 | %81,40 |
| Duyarlılık | %86,70 | %21,20      | %91,90  | %93,10 | %72,30 | %78,30 |
| F1 Skoru   | %86,70 | %16,60      | %91,90  | %93,10 | %71,80 | %79,50 |
| MCC        | %80,30 | %8,50       | %88,10  | %89,80 | %58,80 | %72,50 |

Tablo 3.22. İvmeölçer Veloergometri 8 saniye’de Veloergometri testi esnasında toplanan ivmeölçer verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %88 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 1, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.22.** İvmeölçer Veloergometri 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %88,10 | %20,00      | %92,00  | %92,40 | %72,70 | %79,20 |
| Kesinlik   | %88,10 | %35,90      | %92,10  | %92,60 | %73,10 | %78,80 |
| Duyarlılık | %88,10 | %20,00      | %92,00  | %92,40 | %72,70 | %75,40 |
| F1 Skoru   | %88,10 | %15,80      | %92,00  | %92,30 | %72,00 | %76,50 |
| MCC        | %82,50 | %8,40       | %88,10  | %88,70 | %58,90 | %69,20 |

Tablo 3.23. EKG Veloergometri 2 saniye’de Veloergometri testi esnasında toplanan EKG verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %95 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.23.** EKG Veloergometri 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %88,70 | %22,90      | %97,00  | %97,10 | %83,30 | %88,90 |
| Kesinlik   | %88,60 | %45,10      | %97,00  | %97,10 | %83,40 | %88,70 |
| Duyarlılık | %88,70 | %22,90      | %97,00  | %97,10 | %83,30 | %86,60 |
| F1 Skoru   | %88,60 | %15,90      | %97,00  | %97,10 | %83,20 | %87,40 |
| MCC        | %83,20 | %7,00       | %95,60  | %95,70 | %75,20 | %83,70 |

Tablo 3.24. EKG Veloergometri 4 saniye’de Veloergometri testi esnasında toplanan EKG verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %96 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 3, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.24.** EKG Veloergometri 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %91,60 | %24,00      | %97,90  | %97,70 | %86,00 | %91,70 |
| Kesinlik   | %91,60 | %45,30      | %97,90  | %97,70 | %86,20 | %90,90 |
| Duyarlılık | %91,60 | %24,00      | %97,90  | %97,70 | %86,00 | %90,70 |
| F1 Skoru   | %91,50 | %17,20      | %97,90  | %97,70 | %85,90 | %90,70 |
| MCC        | %87,60 | %10,00      | %96,90  | %96,70 | %79,40 | %87,80 |

Tablo 3.25. EKG Veloergometri 8 saniye’de Veloergometri testi esnasında toplanan EKG verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %97 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 1, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.25.** EKG Veloergometri 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %93,60 | %26,80      | %98,40  | %98,30 | %85,80 | %88,00 |
| Kesinlik   | %93,60 | %49,80      | %98,40  | %98,30 | %85,90 | %87,90 |
| Duyarlılık | %93,60 | %26,80      | %98,40  | %98,30 | %85,80 | %85,50 |
| F1 Skoru   | %93,60 | %20,70      | %98,40  | %98,20 | %85,70 | %86,30 |
| MCC        | %90,60 | %13,60      | %97,70  | %97,40 | %78,90 | %82,30 |

Tablo 3.26. İvmeölçer & EKG Veloergometri 2 saniye’de Veloergometri testi esnasında toplanan ivmeölçer ve EKG verileri ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.26.** İvmeölçer & EKG Veloergometri 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %94,90 | %24,80      | %99,20  | %99,00 | %93,50 | %96,50 |
| Kesinlik   | %94,90 | %36,00      | %99,20  | %99,00 | %93,50 | %96,20 |
| Duyarlılık | %94,90 | %24,80      | %99,20  | %99,00 | %93,50 | %96,10 |
| F1 Skoru   | %94,90 | %17,80      | %99,20  | %99,00 | %93,50 | %96,10 |
| MCC        | %92,50 | %11,00      | %98,80  | %98,60 | %90,40 | %94,80 |

Tablo 3.27. İvmeölçer & EKG Veloergometri 4 saniye'de Veloergometri testi esnasında toplanan ivmeölçer ve EKG verileri ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 1, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.27.** İvmeölçer & EKG Veloergometri 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %96,40 | %25,00      | %99,30  | %99,30 | %92,70 | %96,50 |
| Kesinlik   | %96,40 | %37,50      | %99,30  | %99,30 | %92,80 | %96,20 |
| Duyarlılık | %96,40 | %25,00      | %99,30  | %99,30 | %92,70 | %96,20 |
| F1 Skoru   | %96,40 | %18,10      | %99,30  | %99,30 | %92,70 | %96,20 |
| MCC        | %94,60 | %12,20      | %98,90  | %98,90 | %89,20 | %94,80 |

Tablo 3.28. İvmeölçer & EKG Veloergometri 8 saniye'de Veloergometri testi esnasında toplanan ivmeölçer ve EKG verileri ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 1, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.28.** İvmeölçer & EKG Veloergometri 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %95,80 | %25,80      | %99,40  | %99,20 | %92,80 | %94,50 |
| Kesinlik   | %95,90 | %48,30      | %99,40  | %99,20 | %92,80 | %94,20 |
| Duyarlılık | %95,80 | %25,80      | %99,40  | %99,20 | %92,80 | %94,00 |
| F1 Skoru   | %95,80 | %20,60      | %99,40  | %99,20 | %92,80 | %94,10 |
| MCC        | %93,90 | %13,60      | %99,10  | %98,80 | %89,40 | %91,90 |

**d) GAIN Testi için alınan sonuçlar**

Tablo 3.29. İvmeölçer GAIN 2 saniye'de GAIN ANALYSIS testi esnasında toplanan ivmeölçer verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %86 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 15, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.29.** İvmeölçer GAIN 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %57,50 | %32,60      | %85,10  | %89,80 | %59,00 | %55,90 |
| Kesinlik   | %59,60 | %33,80      | %85,20  | %89,90 | %61,80 | %62,70 |
| Duyarlılık | %57,50 | %32,60      | %85,10  | %89,80 | %59,00 | %52,60 |
| F1 Skoru   | %57,20 | %22,90      | %85,20  | %89,90 | %58,50 | %54,80 |
| MCC        | %40,60 | %11,10      | %79,50  | %86,00 | %43,00 | %38,60 |

Tablo 3.30. İvmeölçer GAIN 4 saniye'de GAIN ANALYSIS testi esnasında toplanan ivmeölçer verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %86 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 9, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.30.** İvmeölçer GAIN 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %72,70 | %32,50      | %86,60  | %90,00 | %68,70 | %72,10 |
| Kesinlik   | %73,20 | %33,80      | %86,60  | %90,10 | %69,70 | %74,30 |
| Duyarlılık | %72,70 | %32,50      | %86,60  | %90,00 | %68,70 | %71,30 |
| F1 Skoru   | %72,70 | %22,60      | %86,60  | %90,00 | %68,70 | %72,50 |
| MCC        | %62,10 | %11,00      | %81,40  | %86,20 | %56,90 | %61,40 |

Tablo 3.31. İvmeölçer GAIN 8 saniye'de GAIN ANALYSIS testi esnasında toplanan ivmeölçer verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %85 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.31.** İvmeölçer GAIN 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %74,80 | %32,70      | %87,70  | %89,60 | %69,70 | %72,70 |
| Kesinlik   | %75,10 | %34,10      | %87,80  | %89,70 | %70,30 | %75,30 |
| Duyarlılık | %74,80 | %32,70      | %87,70  | %89,60 | %69,70 | %71,50 |
| F1 Skoru   | %74,80 | %22,70      | %87,70  | %89,60 | %69,60 | %72,90 |
| MCC        | %65,00 | %11,30      | %83,00  | %85,60 | %58,00 | %62,20 |

Tablo 3.32. EKG GAIN 2 saniye'de GAIN ANALYSIS testi esnasında toplanan EKG verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %97 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 9, mesafe ölçütü ise chebyshev olarak bulunmuştur.

**Tablo 3.32.** EKG GAIN 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %81,30 | %16,00      | %98,20  | %98,30 | %78,10 | %81,30 |
| Kesinlik   | %81,30 | %23,00      | %98,20  | %98,30 | %78,40 | %81,10 |
| Duyarlılık | %81,30 | %16,00      | %98,20  | %98,30 | %78,10 | %80,50 |
| F1 Skoru   | %81,30 | %7,20       | %98,20  | %98,30 | %78,00 | %80,70 |
| MCC        | %74,20 | -%1,00      | %97,50  | %97,60 | %69,90 | %74,20 |

Tablo 3.33. EKG GAIN 4 saniye'de GAIN ANALYSIS testi esnasında toplanan EKG verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.33.** EKG GAIN 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %83,70 | %16,30      | %98,60  | %98,60 | %82,60 | %83,00 |
| Kesinlik   | %83,70 | %28,90      | %98,60  | %98,60 | %82,80 | %83,30 |
| Duyarlılık | %83,70 | %16,30      | %98,60  | %98,60 | %82,60 | %82,80 |
| F1 Skoru   | %83,70 | %7,90       | %98,60  | %98,60 | %82,60 | %83,00 |
| MCC        | %77,60 | -%0,70      | %98,10  | %98,10 | %76,00 | %76,50 |

Tablo 3.34. EKG GAIN 8 saniye'de GAIN ANALYSIS testi esnasında toplanan EKG verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için "k" değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.34.** EKG GAIN 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %86,40 | %17,00      | %99,00  | %98,90 | %85,40 | %85,90 |
| Kesinlik   | %86,40 | %30,70      | %99,00  | %98,90 | %85,60 | %86,70 |
| Duyarlılık | %86,40 | %17,00      | %99,00  | %98,90 | %85,40 | %85,20 |
| F1 Skoru   | %86,40 | %8,40       | %99,00  | %98,90 | %85,40 | %85,80 |
| MCC        | %81,20 | %1,40       | %98,60  | %98,50 | %79,90 | %80,60 |

Tablo 3.35. İvmeölçer ve EKG GAIN 2 saniye’de GAIN ANALYSIS testi esnasında toplanan ivmeölçer ve EKG verileri ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 13, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.35.** İvmeölçer ve EKG GAIN 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %82,20 | %25,30      | %99,40  | %99,40 | %85,10 | %87,90 |
| Kesinlik   | %82,30 | %27,00      | %99,40  | %99,40 | %85,30 | %88,30 |
| Duyarlılık | %82,20 | %25,30      | %99,40  | %99,40 | %85,10 | %87,80 |
| F1 Skoru   | %82,20 | %11,70      | %99,40  | %99,40 | %85,10 | %88,00 |
| MCC        | %75,20 | -%2,70      | %99,20  | %98,50 | %64,40 | %83,30 |

Tablo 3.36. İvmeölçer ve EKG GAIN 4 saniye’de GAIN ANALYSIS testi esnasında toplanan ivmeölçer ve EKG verileri ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 9, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.36.** İvmeölçer ve EKG GAIN 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %90,00 | %25,10      | %99,60  | %99,40 | %89,90 | %91,90 |
| Kesinlik   | %90,00 | %26,50      | %99,60  | %99,40 | %90,10 | %92,00 |
| Duyarlılık | %90,00 | %25,10      | %99,60  | %99,40 | %89,90 | %91,90 |
| F1 Skoru   | %90,00 | %11,70      | %99,60  | %99,40 | %89,90 | %92,00 |
| MCC        | %86,10 | -%3,00      | %99,40  | %99,20 | %86,10 | %88,90 |

Tablo 3.37. İvmeölçer ve EKG GAIN 8 saniye’de GAIN ANALYSIS testi esnasında toplanan ivmeölçer ve EKG verileri ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.37.** İvmeölçer ve EKG GAIN 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %91,00 | %26,40      | %99,60  | %99,50 | %90,50 | %92,40 |
| Kesinlik   | %91,10 | %38,40      | %99,60  | %99,50 | %90,60 | %92,50 |
| Duyarlılık | %91,00 | %26,40      | %99,60  | %99,50 | %90,50 | %92,30 |
| F1 Skoru   | %91,00 | %12,10      | %99,60  | %99,50 | %90,50 | %92,40 |
| MCC        | %87,60 | %2,80       | %99,50  | %99,30 | %86,90 | %89,50 |

#### e) 6MWT testi için alınan sonuçlar

Tablo 3.38. İvmeölçer 6MWT 2 saniye’de 6MWT testi esnasında toplanan ivmeölçer verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %87 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 15, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.38.** İvmeölçer 6MWT 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %62,70 | %37,40      | %86,00  | %90,70 | %61,70 | %65,60 |
| Kesinlik   | %64,30 | %44,70      | %86,30  | %90,90 | %63,90 | %68,10 |
| Duyarlılık | %62,70 | %37,40      | %86,00  | %90,70 | %61,70 | %62,00 |
| F1 Skoru   | %62,00 | %22,50      | %86,00  | %90,70 | %60,70 | %63,70 |
| MCC        | %47,80 | %6,60       | %80,60  | %87,20 | %46,90 | %52,10 |

Tablo 3.39. İvmeölçer 6MWT 4 saniye’de 6MWT testi esnasında toplanan ivmeölçer verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %87 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.39.** İvmeölçer 6MWT 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %76,70 | %37,40      | %87,30  | %91,00 | %67,40 | %76,00 |
| Kesinlik   | %77,00 | %46,20      | %87,50  | %91,20 | %68,30 | %76,80 |
| Duyarlılık | %76,70 | %37,40      | %87,30  | %91,00 | %67,40 | %74,20 |
| F1 Skoru   | %76,60 | %22,20      | %87,30  | %91,00 | %67,10 | %75,20 |
| MCC        | %67,70 | %6,40       | %82,40  | %87,60 | %54,80 | %66,80 |

Tablo 3.40. İvmeölçer 6MWT 8 saniye’de 6MWT testi esnasında toplanan ivmeölçer verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %87 MCC oranı ile RF modeli verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.40.** İvmeölçer 6MWT 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %79,40 | %37,20      | %88,00  | %90,60 | %66,30 | %78,40 |
| Kesinlik   | %79,60 | %43,10      | %88,20  | %90,80 | %67,50 | %79,10 |
| Duyarlılık | %79,40 | %37,20      | %88,00  | %90,60 | %66,30 | %76,80 |
| F1 Skoru   | %79,30 | %21,90      | %88,00  | %90,60 | %65,70 | %77,70 |
| MCC        | %71,40 | %5,70       | %83,40  | %87,00 | %53,20 | %70,20 |

Tablo 3.41. EKG 6MWT 2 saniye’de 6MWT testi esnasında toplanan EKG verisi ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %94 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 9, mesafe ölçütü ise Manhattan olarak bulunmuştur

**Tablo 3.41.** EKG 6MWT 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %77,00 | %22,60      | %96,10  | %96,30 | %74,10 | %70,80 |
| Kesinlik   | %77,00 | %56,60      | %96,10  | %96,30 | %74,90 | %70,40 |
| Duyarlılık | %77,00 | %22,60      | %96,10  | %96,30 | %74,10 | %69,90 |
| F1 Skoru   | %77,00 | %9,00       | %96,10  | %96,30 | %74,00 | %70,10 |
| MCC        | %68,20 | %3,60       | %94,60  | %94,90 | %64,40 | %59,70 |

Tablo 3.42. EKG 6MWT 4 saniye’de 6MWT testi esnasında toplanan EKG verisi ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %96 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.42.** EKG 6MWT 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %79,40 | %22,20      | %97,30  | %97,20 | %75,80 | %70,60 |
| Kesinlik   | %79,40 | %20,90      | %97,30  | %97,20 | %76,50 | %70,80 |
| Duyarlılık | %79,40 | %22,20      | %97,30  | %97,20 | %75,80 | %70,00 |
| F1 Skoru   | %79,40 | %8,20       | %97,30  | %97,20 | %75,70 | %70,20 |
| MCC        | %71,60 | -%0,40      | %96,30  | %96,20 | %66,60 | %59,60 |

Tablo 3.43. EKG 6MWT 8 saniye’de 6MWT testi esnasında toplanan EKG verisi ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %97 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.43.** EKG 6MWT 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %80,40 | %22,20      | %98,10  | %97,90 | %78,80 | %65,70 |
| Kesinlik   | %80,40 | %25,90      | %98,10  | %97,90 | %79,50 | %65,90 |
| Duyarlılık | %80,40 | %22,20      | %98,10  | %97,90 | %78,80 | %64,30 |
| F1 Skoru   | %80,30 | %8,20       | %98,10  | %97,90 | %78,80 | %64,90 |
| MCC        | %72,90 | %0,30       | %97,40  | %97,10 | %70,80 | %52,50 |

Tablo 3.44. İvmeölçer & EKG 6MWT 2 saniye’de 6MWT testi esnasında toplanan ivmeölçer ve EKG verileri ile 2 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %98 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 11, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.44.** İvmeölçer & EKG 6MWT 2 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %83,30 | %39,50      | %99,10  | %99,10 | %82,60 | %87,40 |
| Kesinlik   | %83,50 | %53,50      | %99,10  | %99,10 | %83,00 | %87,20 |
| Duyarlılık | %83,30 | %39,50      | %99,10  | %99,10 | %82,60 | %86,80 |
| F1 Skoru   | %83,20 | %27,60      | %99,10  | %99,10 | %82,60 | %87,00 |
| MCC        | %76,90 | %12,50      | %98,80  | %98,80 | %76,10 | %82,70 |

Tablo 3.45. İvmeölçer & EKG 6MWT 4 saniye’de 6MWT testi esnasında toplanan ivmeölçer ve EKG verileri ile 4 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 7, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.45.** İvmeölçer & EKG 6MWT 4 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %90,90 | %39,50      | %99,40  | %99,30 | %87,50 | %90,90 |
| Kesinlik   | %90,90 | %54,30      | %99,40  | %99,30 | %88,00 | %90,90 |
| Duyarlılık | %90,90 | %39,50      | %99,40  | %99,30 | %87,50 | %90,40 |
| F1 Skoru   | %90,90 | %27,20      | %99,40  | %99,30 | %87,50 | %90,60 |
| MCC        | %87,40 | %12,60      | %99,10  | %99,00 | %83,00 | %87,50 |

Tablo 3.46. İvmeölçer & EKG 6MWT 8 saniye’de 6MWT testi esnasında toplanan ivmeölçer ve EKG verileri ile 8 saniyelik zaman aralığında oluşturulmuş veri seti için sonuçlar görülmektedir. En iyi sonucu %99 MCC oranı ile XGBoost ve RF modelleri verirken en kötü sonuç ise Naive Bayes modelinde olduğu görülmektedir. Grid Search ile yapılan aramada KNN için “k” değeri 5, mesafe ölçütü ise Manhattan olarak bulunmuştur.

**Tablo 3.46.** İvmeölçer & EKG 6MWT 8 saniye

| Metrikler  | KNN    | Naive Bayes | XGBoost | RF     | MLP    | 1D-CNN |
|------------|--------|-------------|---------|--------|--------|--------|
| Doğruluk   | %92,10 | %39,10      | %99,50  | %99,30 | %85,80 | %92,10 |
| Kesinlik   | %92,10 | %51,10      | %99,50  | %99,30 | %86,20 | %91,90 |
| Duyarlılık | %92,10 | %39,10      | %99,50  | %99,30 | %85,80 | %91,90 |
| F1 Skoru   | %92,00 | %26,40      | %99,50  | %99,30 | %85,80 | %91,90 |
| MCC        | %89,00 | %11,70      | %99,30  | %99,10 | %80,50 | %89,20 |



## 4. TARTIŞMA

Bu bölümde ilgili testler (Stair, Tug vb.) özelinde sensör bazlı (ivmeölçer gibi) 2, 4 ve 8 saniyelik çalışmalar ayrı ayrı incelenmiş sonrasında zaman ölçeği ele alınmıştır. Tüm sensörler için bu işlem tekrarlandıktan sonra İvmeölçer, EKG veya bu iki sensörün birlikte kullanımının hangisinin daha verimli olduğu incelenecektir. Yapılan çalışmada MCC performans metriği diğer metrikleri de kapsaması ve doğruluk gibi bir başarı yanığı yaratmaması nedeniyle ana değerlendirme metriği olarak seçilmiş ve başarı kriteri olarak 0,750 seviyesi seçilmiştir. Bu seviyenin altında olan tüm modeller başarısız üstünde olan modeller ise başarılı kabul edilecektir.

Sonuç kısmında verilen tablolara genel bir bakış atıldığında ilk göze çarpan sonuç, Naive Bayes modelinin anlamlı bir başarıya ulaşamamış olmasıdır. Bunun nedeni problemin çok boyutlu veri kümesi nedeniyle karmaşık bir problem olması aynı zamanda algoritmanın temel mantığında olan tüm özellikler arası bağımsızlık varsayımının bu problem için uygun olmamasıdır. Bu nedenle yapılan incelemelerde Naive Bayes dikkate alınmamıştır.

### a) Stair

Tablo 3.2.'de verilen performans başarı metrikleri sonuçlarına göre, doğruluk baz alındığında, en yüksek sonuçlar XGBoost ve RF modellerinden alınmıştır. Birbirlerine oldukça yakın sonuçlar veren bu iki model kırılabilirliği neredeyse kesin bir sonuç ile tahmin edebildiği tespit edilmiştir. Ancak diğer metrik sonuçlarına bakıldığında ise en başarılı modelinin RF olduğu görülmektedir. XGBoost'da ise özellikle MCC değerine bakıldığında göz ardı edilebilecek bir seviyede düşüş görülmektedir. Bunun nedeni düzenli bir dağılıma sahip olmayan veri setleri sonucunda oluşan doğruluk yanığıdır. Bunun yanı sıra, KNN, MLP, 1D-CNN modellerinin doğruluk ölçümlerinin 0,8'e yakınsadığı görülmektedir. Diğer metriklerle birlikte okunduğunda ise bu üç metrik arasında en başarılıdan başarısız doğru sırayla 1D-CNN, KNN ve MLP olduğu söylenebilir. MCC değerlerinin üç model içinde 0.75 değerinin altında olması göz önüne alındığında kullanım için uygun değildir. Bu incelemelerin sonucunda ivmeölçer verilerinden oluşan 2 saniye ölçekli Stair testi esnasında en başarılı tahminlemenin sırasıyla RF, XGBoost olduğu bu modellerin yüksek doğruluk ile tahmin yaptığı diğer modellerin sonuçlarının ise güvenilirmez olduğu yargısına ulaşılabilir.

Tablo 3.3.'te verilen performans başarı metrikleri sonuçlarına göre, en yüksek sonuçlar XGBoost ve RF modellerinden alınmıştır. Birbirlerine oldukça yakın sonuçlar veren bu iki model kırılabilirliği neredeyse kesin bir sonuç ile tahmin edebildiği görülmektedir. KNN için MCC değerindeki yaklaşık 0,1'lik düşüş olmasına karşın 0,75'lik başarı oranını geçememiş ve başarısız olmuştur. 1D-CNN ve MLP ise istenilen seviyede başarı gösterememiştir. Bu incelemelerin sonucunda ivmeölçer verilerinden oluşan 4 saniye ölçekli Stair testi esnasında en başarılı tahminlemenin sırasıyla RF, XGBoost olduğu bu modellerin yüksek doğruluk ile tahmin yaptığı diğer modellerin sonuçlarının ise güvenilmez olduğu yargısına ulaşılabilir.

Tablo 3.4.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı model XGBoost modeli olmuştur. RF başarı metrikleri yine XGBoost ile yakınsamasına rağmen ivmeölçer verilerinde XGBoost'un gerisinde kalmıştır. XGBoost ile RF haricindeki diğer modellerin istenilen başarıyı gösteremediği değerlendirilmiştir. Bu incelemelerin sonucunda ivmeölçer verilerinden oluşan 8 saniye ölçekli Stair testi esnasında en başarılı tahminlemenin sırasıyla RF, XGBoost olduğu bu modellerin yüksek doğruluk ile tahmin yaptığı diğer modellerin sonuçlarının ise güvenilmez olduğu yargısına ulaşılabilir.

Tablo 3.2., 3.3. ve 3.4. sonuçları incelendiğinde, 2 saniyelik ve 4 saniyelik ölçeklendirmede RF başarı metriklerinde bir değişim olmadığı, XGBoost modeli için başarının arttığı görülmektedir. Ölçeklendirme 4 saniyeden 8 saniyeye çıkarıldığında ise XGBoost haricindeki tüm modellerin başarılarının düştüğü, XGBoost modelinde ise anlamlı olmayan minimal bir artış olduğu belirlenmiştir. MLP ve 1D-CNN için ise 2 saniyeden 8 saniyeye çıktıkça düşen bir başarı görülmektedir. Bu inceleme sonucunda ivmeölçer verileri ile Stair testinin en iyi sonuç verdiği ölçeklendirme süresinin 4 saniye olduğu görülmektedir.

Tablo 3.5.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı model XGBoost ve RF modelleridir. Bu iki model arasında kayda değer bir fark görülememekte ve neredeyse kesin doğru sonuçlarla tahminleme yapabilmektedir. Yüksek doğruluk oranıyla tahminleme yapan 1D-CNN modeli kırılabilirliğin belirlenmesinde kullanılabilir uygun bir model olarak öne çıkmaktadır. Sırasıyla KNN ve MLP modellerinin de doğru tahminlemelerde bulunduğu söylenebilir. İncelemenin sonunda EKG verilerinden oluşan 2 saniye ölçekli Stair testinde Naive bayes harici tüm modellerin başarılı tahminleme yaptığı ancak özellikle XGBoost ve RF modellerinin neredeyse kesin doğrulukta tahminleme yaptığı sonucuna ulaşılmıştır.

Tablo 3.6.'da verilen performans başarı metrikleri sonuçlarına göre, RF ve XGBoost aynı oranlarda başarı göstermekte ve neredeyse kesin doğrulukta tahminleme yapmaktadır. KNN ve 1D-CNN de aynı şekilde yüksek oranda doğru tahminleme yapabilmektedir. MLP modeli ise yine yüksek doğruluk oranına sahip ve gerçek uygulamalarda kullanılabilir bir model olarak görünse de diğer modellerin arkasında kalmıştır. Sonuç olarak EKG verilerinden oluşan 4 saniye ölçekli Stair testinde Naive Bayes haricinde tüm modeller başarılı tahminlemeler yapmış ancak bu başarılı modeller arasında yine en çok göze çarpan modeller XGBoost ve RF modelleri olmuştur.

Tablo 3.7.'de verilen performans başarı metrikleri sonuçlarına göre, RF ve XGBoost algoritmaları yaklaşık 0,98 oranında kesine yakın doğrulukta tahminlemeler yaptığı görülmektedir. Sırasıyla KNN, 1D-CNN ve MLP en başarılı algoritmalarıdır. Naive Bayes haricindeki tüm modellerin başarı metrikleri 0,80 seviyesinin üzerinde kalmıştır. Sonuç olarak Naive Bayes hariç tüm modellerin başarıya ulaştığı yargısına varılmıştır.

Tablo 3.5., 3.6. ve 3.7. sonuçları incelendiğinde, kullanılan modellerin başarı metriklerinin oldukça yüksek olduğu, RF ve XGBoost modellerinin 1'e yakınsadığı görülmektedir. EKG verileri kullanılarak yapılan ölçümlerde 2 saniyeden 4 saniyeye çıkmasında 1D-CNN haricinde tüm modellerin başarı oranları artmıştır. Ancak bu artışlar KNN haricinde minimal artışlardır. 4 saniyeden 8 saniyeye çıkışta ise RF, XGBoost modellerinin metrikleri aynı kalmış, KNN ve MLP modellerindeki başarı metrikleri minimal bir artış göstermiş, 1D-CNN modelinin başarı metrikleri ise düşmüştür. Stair testi esnasında EKG verileri ile oluşturulan veri setlerinde genel olarak modellerin başarıları göz önüne alındığında optimum ölçeklendirme olarak 4 saniye görülmektedir. Bunun nedeni 2 saniye ve 4 saniye arasında modellerdeki genel artış eğiliminin 4 saniyeden 8 saniyeye çıkışta görülememesidir.

Tablo 3.8.'de verilen performans başarı metrikleri sonuçlarına göre, binde sekizlik bir hata payı ile XGBoost ve binde dokuzluk bir hata payı ile RF modellerinin tahminleme yapması dikkat çekicidir. Bununla birlikte sırasıyla 1D-CNN, KNN ve MLP modelleri de kesine yakın doğrulukta tahminleme yaptığı görülmektedir. Sonuç olarak 2 saniye ölçeğinde Stair testi esnasında toplanan ivmeölçer ve EKG verilerinin birlikte kullanılması ile Naive Bayes haricindeki tüm modellerin başarılı sonuç verdiği tespit edilmiştir.

Tablo 3.9.'da verilen performans başarı metrikleri sonuçlarına göre, binde yedilik bir hata payı ile XGBoost ve binde sekizlik bir hata payı ile RF modellerinin tahminleme yapması dikkat çekicidir. KNN ve 1D-CNN yüksek bir doğruluk oranı ile tahminleme yaparken MLP modeli ise kabul edilebilir doğrulukta tahminleme yapmaktadır. Sonuç olarak 4 saniye ölçeğinde Stair testi esnasında toplanan ivmeölçer ve EKG verilerinin birlikte kullanılması ile Naive Bayes haricindeki tüm modellerin başarılı sonuç verdiği görülmüştür.

Tablo 3.10.'da verilen performans başarı metrikleri sonuçlarına göre, binde yedilik bir hata payı ile XGBoost ve binde sekizlik bir hata payı ile RF modellerinin tahminleme yapması dikkat çekicidir. KNN modelinin ise MCC metriği 0,90'ın üzerindedir. 1D-CNN ve MLP modellerinin ise yüksek isabet oranı ile tahminleme yaptığı görülmektedir. Sonuç olarak 8 saniye ölçeğinde Stair testi esnasında toplanan ivmeölçer ve EKG verilerinin birlikte kullanılması ile Naive Bayes haricindeki tüm modellerin başarılı sonuç verdiği değerlendirilmiştir.

Tablo 3.8., 3.9. ve 3.10. sonuçları incelendiğinde, ölçeklendirme 2 saniyeden 4 saniyeye çıkarıldığında RF ve XGboost haricindeki modellerin MCC metriklerinde düşüş gözlemlenmektedir. Sonucun doğruluğunu olumsuz yönde etkilediği söylenebilir. Ancak 4 saniyeden 8 saniyeye veya 2 saniyeden 8 saniyeye çıktığımızda MCC metriğinin arttığı gözlemlenmiştir. Fakat bu artış kayda değer büyüklükte bir artış değildir. Bu nedenle 2 saniyelik ölçeklendirme, Stair testi esnasında toplanan ivmeölçer ve EKG verilerinin birlikte kullanılmasından oluşan veri setinde en iyi sonucu vermiştir.

Stair testi esnasında kırılğanlığın değerlendirilmesi konusunda en başarılı modeller hiç şüphesiz ki XGBoost ve RF modelleri olmuştur. Bununla birlikte 2 saniyelik ivmeölçer ve EKG verilerinin birlikte kullanılması ile oluşturulan veri seti ile kırılğanlığın tespitinde en başarılı sonuçlar alınmıştır. Bu sonuç göstermektedir ki merdiven inip çıkarken İvmeölçer ve kalp ritminin birlikte değerlendirilmesi kırılğanlığın doğru tahminlenme oranını kayda değer oranda arttırmaktadır. İlginç olarak sadece EKG verilerinin kullanılarak kırılğanlık tahminleme yapılan sonuçlar ivmeölçer verilerinin kullanılarak yapılan tahminlerden daha isabetli olduğu anlaşılmıştır.

## b) TUG

Tablo 3.11.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı model yüksek doğrulukta tahminleme yapan RF modelidir. XGBoost modeli ise RF modeline yakın sonuçlar vermiştir. Ancak sırasıyla 1D-CNN, MLP ve KNN başarı metriklerinde istenilen sonuca ulaşamamıştır. Tüm bu veriler incelendiğinde TUG testi 2 saniyelik zaman ölçeğinde İvmeölçer verileriyle yapılan tahminlemede RF ve XGBoost modellerinin başarılı diğer modellerin ise istenilen başarıya ulaşamadığı görülmüştür.

Tablo 3.12.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı model yüksek doğrulukta tahminleme yapan RF modelidir. XGBoost modeli ise RF modeline yakın sonuçlar vermiştir. KNN modeli için MCC değerine bakıldığında istediğimiz başarı kriteri olan 0,75 değerine yakınsadığı ancak altında kaldığı görülmektedir. 1D-CNN ve MLP ise istenilen başarıya ulaşamamıştır. Sonuç olarak TUG testi 4 saniyelik zaman ölçeğinde İvmeölçer verileriyle yapılan tahminlemede RF ve XGBoost modellerinin başarılı olduğu, diğer modellerin ise istenilen başarıya ulaşamadığı görülmektedir.

Tablo 3.13.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller başarı metrikleri birbirine yakınsadığı görülen RF ve XGBoost modelleridir. KNN modeli için MCC değerine bakıldığında istediğimiz başarı kriteri olan 0,75 değerine yakınsadığı ancak altında kaldığı görülmektedir. Sonuç olarak TUG testi 8 saniyelik zaman ölçeğinde İvmeölçer verileriyle yapılan tahminlemede RF ve XGBoost modellerinin başarılı olduğu, diğer modellerin ise istenilen başarıya ulaşamadığı değerlendirilmektedir.

Tablo 3.11., 3.12. ve 3.13. sonuçları incelendiğinde 2 saniyelik ölçekten 4 saniyeye çıkıldığında KNN modelindeki MCC değerinde artış görülmekte ancak bu istenen 0,75 eşliğini geçememektedir. XGBoost modelinde minimal bir artış olmuş ancak RF, MLP ve 1D-CNN MCC değerlerinde düşüş görülmüştür. 4 saniyelik ölçek 8 saniyeye çıkarıldığında ise KNN MCC değeri aynı kalırken diğer tüm modellerin MCC değerinde düşüş meydana geldiği saptanmıştır. Sonuç olarak TUG testi esnasında toplanan ivmeölçer verilerinin analizinin en doğru sonucu verdiği zaman ölçeği 2 saniyelik zaman ölçeği olarak tespit edilmiştir.

Tablo 3.14.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller birbirine çok yakın sonuçlar veren RF ve XGBoost modelleridir. Bu iki modelin

MCC değerine bakıldığında neredeyse kesin doğrulukta bir tahminleme yaptığı görülmektedir. Ayrıca diğer modellerden ayrılarak 0,9 MCC değerinin üzerinde sonuç veren 1D-CNN modeli göze çarpmaktadır. Aynı şekilde Sırasıyla KNN ve MLP modellerinin de yüksek MCC değerlerine sahip olduğu ve başarılı bir şekilde tahminleme yapabildiği görülmektedir. Sonuç olarak 2 saniye ölçeğinde TUG testi esnasında EKG verilerinin analizi ile Naive Bayes haricindeki tüm modellerin başarılı sonuç verdiği görülmektedir.

Tablo 3.15.'te verilen performans başarı metrikleri sonuçlarına göre, birbirine yakınsayan MCC değerleriyle en doğru tahminlemeyi yapan modellerin RF ve XGBoost olduğu değerlendirilmektedir. KNN modeli 0,75'lik başarı sınırı olarak belirlenen değere yaklaşıp da altında kalmıştır. MLP ve 1D-CNN modelleri ise istenilen başarı seviyesine ulaşamamıştır. Sonuç olarak 4 saniye ölçeğinde TUG testi esnasında toplanan EKG verilerinin analizi ile RF ve XGBoost modellerinin başarılı tahminleme yaptığı tespit edilmiştir.

Tablo 3.16.'da verilen performans başarı metrikleri sonuçlarına göre, en başarılı model XGBoost ve RF modelleridir. Bu iki modelin MCC değerine bakıldığında neredeyse kesin doğrulukta bir tahminleme yapmaktadır. KNN modeli ise MCC değeri alt başarı eşiğine yakın olmasıyla birlikte başarılı bir tahminleme yapabildiği görülmektedir. Ancak MLP ve 1D-CNN modelleri başarılı bir tahminleme yapamamıştır. Sonuç olarak 8 saniye ölçeğinde TUG testi esnasında toplanan EKG verilerinin analizi ile RF, XGBoost ve KNN modellerinin başarılı tahminleme yaptığı görülmüştür.

Tablo 3.14., 3.15. ve 3.16. sonuçları incelendiğinde, genel olarak 2 saniyelik ölçekten 8 saniyelik ölçeğe gidildiğinde tahminleme başarısının düştüğü gözlenmiştir. Ancak bu düşüş 2 saniyelik ölçekten 4 saniyelik ölçeğe geçerken dramatik bir düşüş 4 saniyelik ölçekten 8 saniyeye çıkarken bir toparlanma şeklindedir. Bu nedenle TUG testi esnasında EKG verilerinin kullanılarak yapılan kırılma testi için en iyi zaman ölçeğinin 2 saniyelik zaman dilimi olduğu yargısına varılmıştır.

Tablo 3.17.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller birbirine çok yakın sonuçlar veren RF ve XGBoost modelleridir. Bu iki modelin MCC değerine bakıldığında neredeyse kesin doğrulukta bir tahminleme yaptığı görülmektedir. Ayrıca diğer modellerden ayrılarak 0,9 MCC değerinin üzerinde sonuç veren 1D-CNN modeli göze çarpmaktadır. Aynı şekilde Sırasıyla KNN ve MLP modellerinin de

yüksek MCC değerlerine sahip olduğu ve başarılı bir şekilde tahminleme yapabildiği görülmektedir. Sonuç olarak 2 saniye ölçeğinde TUG testi esnasında ivmeölçer ve EKG verilerinin analizi ile Naive Bayes haricindeki tüm modellerin başarılı sonuç verdiği değerlendirilmiştir.

Tablo 3.18.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller birbirine çok yakın sonuçlar veren RF ve XGBoost modelleridir. Bu iki modelin MCC değerine bakıldığında neredeyse kesin doğrulukta bir tahminleme yaptığı görülmektedir. Sırasıyla 1D-CNN, KNN ve MLP modelleri 0,8 üzerindeki MCC değerleriyle oldukça başarılı tahminlemeler yaptığı görülmüştür. Sonuç olarak 4 saniye ölçeğinde TUG testi esnasında ivmeölçer ve EKG verilerinin analizi ile Naive Bayes haricindeki tüm modellerin başarılı sonuç verdiği tespit edilmiştir.

Tablo 3.19.'da verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller birbirine çok yakın sonuçlar veren RF ve XGBoost modelleridir. Bu iki modelin MCC değerine bakıldığında neredeyse kesin doğrulukta bir tahminleme yaptığı görülmektedir. Ayrıca diğer modellerden ayrılarak 0,9 MCC değerinin üzerinde sonuç veren KNN modeli göze çarpmaktadır. KNN modeli ile 1D-CNN ve MLP modellerinin MCC değerlerine bakıldığında yüksek doğruluk ile tahminleme yapabildiği görülmektedir. Sonuç olarak, 4 saniye ölçeğinde TUG testi esnasında ivmeölçer ve EKG verilerinin analizi ile Naive Bayes haricindeki tüm modellerin başarılı sonuç verdiği görülmüştür.

Tablo 3.17., 3.18. ve 3.19. sonuçları incelendiğinde, zaman ölçekleri ile başarı metrikleri arasında anlamlı bir ilişki bulunmamaktadır. Açıklamak gerekirse, ölçek farklılıkları önceki örneklerde olduğu gibi bir değişime sebebiyet vermektedir. Ancak bu değişiklikler (düşüş/artış) oldukça sınırlı kalmıştır. Örneğin 2 saniyelik ölçekten 8 saniyelik ölçeğe çıkıldığında KNN modeli için MCC değerinde 0,023'lük bir artış yaşanırken 1D-CNN modelinde 0,013'lük bir azalış görülmektedir. Sonuç olarak, TUG testi esnasında ivmeölçer verileri ve EKG verilerinin birlikte analizi ile 3 modelin MCC değerlerinin 0,9 bandını aştığı 2 saniyelik zaman ölçeği daha makul görülmektedir.

TUG testi esnasında kullanılan veri yapılarında en başarılı sonucu veren ivmeölçer ve EKG'nin birlikte kullanıldığı veri setleridir. Aynı zamanda EKG verilerinin yalnız kullanıldığı veri setlerinde de başarı oranlarının büyük oranda arttığı görülmektedir. İvmeölçer verilerinde ise RF ve XGBoost modelleri haricindeki tüm modellerin başarılı

tahminleme yapmada zorlandığı görülmektedir. Bununla birlikte RF ve XGBoost modellerinin de MCC değerleri sadece ivmeölçer verilerinin kullanıldığı veri setini işlerken düşüş göstermesi dikkat çekicidir.

### c) Veloergometri

Tablo 3.20.'de verilen performans başarı metrikleri sonuçlarına göre, en iyi MCC değerini RF modelinin verdiği görülmektedir. RF modeli kesine yakın doğrulukta tahminleme yapmaktadır. XGBoost yüksek oranda başarılı tahminlemeler yaparken, KNN ve 1D-CNN modelleri için başarılı olarak nitelendirilebilecek MCC değerlerine ulaştığı görülmüştür. MLP ise başarılı bir sonuç verememiştir. Sonuç olarak MLP ve Naive Bayes modelleri haricinde tüm modeller başarılı tahminlemeler yapabilmektedir.

Tablo 3.21.'de verilen performans başarı metrikleri sonuçlarına göre, sırasıyla RF, XGBoost ve KNN modelleri başarılı tahminlemeler yapabilmıştır. 1D-CNN modeli ise iyi tahminleme yapabildiği ancak başta da verilen 0,75 MCC başarı kriterine ulaşamadığı için başarısız kabul edilmektedir. MLP modeli ise başarı kriterinin çok altında kalmıştır. Sonuç olarak RF, XGBoost ve KNN modelleri başarılı tahminlemeler yaparken MLP ve Naive Bayes modelleri başarısız olmuştur.

Tablo 3.22.'de verilen performans başarı metrikleri sonuçlarına göre, RF, XGBoost ve KNN modelleri başarılı tahminlemeler yapabilmıştır. Ancak buradaki dikkat çekici husus KNN modelinde “K” değerinin bir çıkmasıdır.  $K=1$  olduğunda genelleme kabiliyetinin olmadığı ve aşırı öğrenmeye müsait bir yapı oluşturduğunu bu nedenle ise modelin güvenilir olmadığını söylenebilir. MLP ve 1D-CNN modelleri ise hedeflenen MCC değerinden çok uzakta kalarak başarısız olmuştur. Sonuç olarak RF, XGBoost modelleri başarılı tahminlemeler yaparken MLP ve Naive Bayes modelleri başarısız olmuştur. KNN modeli ise güvenilir olarak göze çarpmaktadır.

Tablo 3.20., 3.21. ve 3.22. sonuçları incelendiğinde, 2 saniyelik ölçekten 4 saniyelik ölçeğe çıkıldığında KNN ve XGBoost başarı metriklerinin artış gösterdiği görülmektedir. RF ve 1D-CNN modellerinin başarı metrikleri ise aksine düşüş göstermiştir. Modeller üzerinde yaşanan düşüş ve artışlar karşılaştırıldığında 4 saniyelik ölçeklendirmenin 2 saniyelik ölçeklendirmeden daha verimli olduğu sonucuna varılmıştır. 4 saniyelik ölçeklendirmeden 8 saniyeye çıkıldığında ise XGBoost modeli sabit kalmış, KNN modeli

ise güvenilir hale gelmiştir. Diğer modellerin başarı metrikleri ise düşüş göstermiştir. Sonuç olarak, Veloergometri testi esnasında toplanan ivmeölçer verilerinin analizi ile yapılan kırılma tahminlemesi için en uygun zaman ölçeği 4 saniye olarak tespit edilmiştir.

Tablo 3.23.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modellerin RF ve XGBoost modelleri olduğu görülmektedir. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. Sırasıyla 1D-CNN ve KNN modellerinin ise yüksek doğrulukta tahminleme yaparak başarılı olduğu görülmüştür. MLP modeli ise başarı kriteri olarak belirlenen MCC değerinin üzerinde bir sonuç alarak başarılı olmuştur. Sonuç olarak Naive Bayes modeli haricindeki tüm modellerin başarılı bir şekilde tahminleme yapmıştır.

Tablo 3.24.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modellerin XGBoost ve RF modelleri olduğu görülmektedir. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. Sırasıyla 1D-CNN, KNN ve MLP modellerinin ise yüksek doğrulukta tahminleme yaparak başarılı olduğu görülmektedir. Sonuç olarak Naive Bayes modeli haricindeki tüm modellerin başarılı bir şekilde tahminleme yapmıştır.

Tablo 3.25.'te verilen performans başarı metrikleri sonuçlarına göre en başarılı modellerin XGBoost ve RF modelleri olduğu görülmektedir. Ancak buradaki dikkat çekici husus KNN modelinde “K” değerinin bir çıkmasıdır. K=1 olduğunda genelleme kabiliyetinin olmadığı ve overfit (aşırı öğrenme) olmaya müsait bir yapı oluşturduğunu bu nedenle ise modelin güvenilir olmadığını söylenebilir. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. Sırasıyla 1D-CNN modelinin yüksek doğrulukta tahminleme yaparak başarılı olurken, MLP modeli ise başarı kriteri olarak belirlenen MCC değerinin üzerinde bir sonuç alarak başarılı olmuştur. Sonuç olarak XGBoost, RF, 1D-CNN ve MLP modellerinin başarılı tahminlemeler yaptığı tespit edilmiştir.

Tablo 3.23., 3.24. ve 3.25. sonuçları incelendiğinde, 2 saniyelik ölçekten 4 saniyelik ölçeğe çıkıldığında tüm modellerin başarı metriklerinde anlamlı yükselişler gözlenmiş ve özellikle MLP modeli daha güvenilir tahminleme yapabilir hale gelmiştir. 4 saniyelik ölçek 8 saniyeye çıkarıldığında ise KNN modeli güvenilir hale gelirken, diğer modellerde anlamlı bir yükseliş olmamıştır. Sonuç olarak, Veloergometri testi esnasında toplanan EKG verilerinin analizi ile yapılan kırılma tahminlemesinde en uygun zaman ölçeği 4 saniye olarak tespit edilmiştir.

Tablo 3.26.'da verilen performans başarı metrikleri sonuçlarına göre, tüm modellerin oldukça başarılı olduğu görülmektedir. Sırasıyla XGBoost, RF, 1D-CNN, KNN ve MLP modellerinin hepsi kesin sonuca yakın olarak tahminleme yapmaktadır. Sonuç olarak Naive Bayes haricindeki tüm modellerin yüksek oranda başarılı tahminleme yapabilmektedir.

Tablo 3.27.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modellerin XGBoost, RF ve 1D-CNN olduğu görülmektedir. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. KNN modelinin başarı metrikleri yüksek olmasına rağmen "K" değeri göz önünde bulundurulduğunda model güvenilir değildir. MLP modeli ise yüksek doğruluk oranı ile tahminleme yaparak başarılı olmuştur. Sonuç olarak, KNN modelinin güvenilir olmadığı ve XGBoost, RF, 1D-CNN ve MLP modellerinin başarılı olduğu değerlendirilmektedir.

Tablo 3.28.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modellerin XGBoost, RF ve 1D-CNN olduğu görülmektedir. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. KNN modelinin başarı metrikleri yüksek olmasına rağmen "K" değeri göz önünde bulundurulduğunda model güvenilir değildir. MLP modeli ise yüksek doğruluk oranı ile tahminleme yaparak başarılı olmuştur. Sonuç olarak, KNN modelinin güvenilir olmadığı ve XGBoost, RF, 1D-CNN ve MLP modellerinin başarılı olduğu saptanmıştır.

Tablo 3.26., 3.27. ve 3.28. sonuçları incelendiğinde, 2 saniyelik ölçekten 4 saniyelik ölçeye çıkıldığında hiçbir modelin performansında anlamlı bir yükseliş görülmezken KNN modelinin güvenilirliğini yitirdiği belirlenmiştir. 4 saniyelik ölçekten 8 saniyelik ölçeye çıkıldığında ise modellerin performansında anlamlı bir yükseliş yaşanmamış ve KNN güvenilir durumunu korumuştur. Sonuç olarak, Veloergometri testi esnasında toplanan ivmeölçer ve EKG verilerinin birlikte analiz edilerek Kırılma tahminlemesinde en uygun zaman ölçeği 2 saniye olarak tespit edilmiştir.

Veloergometri testi esnasında kullanılan veri yapılarında en başarılı sonucu veren ivmeölçer ve EKG'nin birlikte kullanıldığı veri setleridir. Aynı zamanda EKG verilerinin yalnız kullanıldığı veri setlerinde de başarı oranlarının büyük oranda arttığı görülmektedir. İvmeölçer verilerinde ise RF ve XGBoost modelleri haricindeki tüm modellerin başarılı tahminleme yapmada zorlandığı görülmektedir. Bununla birlikte RF ve XGBoost

modellerinin de MCC deęerleri sadece ivmeölçer verilerinin kullanıldığı veri setini işlerken düşüş göstermesi dikkat çekicidir.

#### **d) Gaint**

Tablo 3.29.'da verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin RF olduğu görülmektedir. RF modeli yüksek başarı oranıyla tahminleme yapabilmektedir. 0,79 MCC deęeriyle XGBoost modelinin de başarılı tahminleme yapabildięi görülmektedir. KNN modelinde Grid Search ile belirlenen "k" parametresi verilen potansiyel "k" kümesinin son elemanı olan 15 olarak belirlendięi göz önünde bulundurulduğunda performanstaki bu düşüklüğün nedeninin uygun "k" parametresinin belirlenememesi olduğu ifade edilebilir. MLP ve 1D-CNN modelleri ise 0,75 MCC başarı kriterinden oldukça düşük sonuçlar vererek başarısız olmuşlardır. Sonuç olarak, 2 saniye ölçeğinde Gaint Analysis testi esnasında toplanan ivmeölçer verilerinin analiz edilerek Kırılgenlik tahminlemede RF ve XGBoost modellerinin başarılı tahminleme yapabildięi ortaya çıkmıştır.

Tablo 3.30.'da verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin RF olduğu görülmektedir. RF ile XGBoost modelleri yüksek doğruluk oranıyla tahminleme yapmaktadır. KNN, MLP ve 1D-CNN modelleri ise 0,75 MCC başarı kriterinin oldukça altında sonuç vererek başarısız olmuşlardır. Sonuç olarak, 4 saniye ölçeğinde Gaint Analysis testi esnasında toplanan ivmeölçer verilerinin analiz edilerek Kırılgenlik tahminlemede RF ve XGBoost modellerinin başarılı tahminleme yapabildięi görülmektedir.

Tablo 3.31.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin RF olduğu görülmektedir. RF ile XGBoost modelleri yüksek doğruluk oranıyla tahminleme yapmaktadır. KNN, MLP ve 1D-CNN modelleri ise 0,75 MCC başarı kriterinin oldukça altında sonuç vererek başarısız olmuşlardır. Sonuç olarak, 8 saniye ölçeğinde Gaint Analysis testi esnasında toplanan ivmeölçer verilerinin analiz edilerek Kırılgenlik tahminlemede RF ve XGBoost modellerinin başarılı tahminleme yapabildięi görülmüştür.

Tablo 3.29., 3.30. ve 3.31. sonuçları incelendiğinde, 2 saniyelik ölçekten 4 saniyelik ölçeğe çıkıldığında XGBoost modelinin performansında minimal bir yükseliş ile 0,8 MCC

değerini geçtiği görülmektedir. KNN modelinde ise “K” değerinin verilen “K” kümesi içerisinde bir değere geldiği ancak başarı metriklerinde yeterli bir yükseliş görülemediği. 4 saniyelik ölçek 8 saniyeye çıkarıldığında ise hiçbir anlamlı yükseliş görülemediği gibi RF modelinin başarı metriklerinde minimal bir düşüş gözlemlenmiştir. Sonuç olarak, kayda değer bir etki görülmemekle birlikte Gaint Analysis testi esnasında toplanan ivmeölçer verilerinin analiz edilerek Kırılgnlık tahminlemesinde en uygun zaman ölçeği 4 saniye olarak belirlenmiştir.

Tablo 3.32.’te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin birbirine yakın sonuçlarla RF ve XGBoost olduğu saptanmıştır. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. KNN, 1D-CNN ve MLP modelleri ise başarı kriteri olan 0,75 MCC değerini geçememiş ve başarısız olmuştur. Sonuç olarak, 2 saniye ölçeğinde Gaint Analysis testi esnasında toplanan EKG verilerinin analiz edilerek Kırılgnlık tahminlemesinde RF ve XGBoost modellerinin başarılı tahminleme yapabildiği tespit edilmiştir.

Tablo 3.33.’te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin birbirine yakın sonuçlarla RF ve XGBoost olduğu saptanmıştır. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. KNN, 1D-CNN ve MLP modelleri ise başarı kriteri olan 0,75 MCC değerini geçerek başarılı olmuştur. Sonuç olarak, 4 saniye ölçeğinde Gaint Analysis testi esnasında toplanan EKG verilerinin analiz edilerek Kırılgnlık tahminlemesinde Naive Bayes haricindeki tüm modellerin başarılı tahminleme yapabildiği tespit edilmiştir.

Tablo 3.34.’te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin birbirine yakın sonuçlarla RF ve XGBoost olduğu görülmüştür. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. Sırasıyla KNN, 1D-CNN ve MLP modelleri ise yüksek doğruluk oranlarında başarılı tahminleme yapabildiği görülmektedir. Sonuç olarak, 8 saniye ölçeğinde Gaint Analysis testi esnasında toplanan EKG verilerinin analiz edilerek Kırılgnlık tahminlemesinde Naive Bayes haricindeki tüm modellerin başarılı tahminleme yapabildiği saptanmıştır.

Tablo 3.32., 3.33. ve 3.34. sonuçları incelendiğinde, 2 saniyelik ölçekten 4 saniyelik ölçeğe çıkıldığında sınır altında kalmış olan KNN, MLP ve 1D-CNN modellerinin başarılı tahminleme yapabilir hale geldiği ve aynı zamanda RF ve XGBoost modellerinin ise tahmin

doğruluğunun daha fazla arttığı tespit edilmiştir. 4 saniyelik ölçekten 8 saniyelik ölçeğe çıkıldığında ise KNN, MLP ve 1D-CNN modellerinin 0,8 MCC değerini aşarak (MLP 0,799) tahminlemelerini iyileştirmiştir. Sonuç olarak, Gaint Analysis testi esnasında toplanan EKG verileri analiz edilerek Kırılgnlık tahminlemesinde en uygun zaman ölçeği 8 saniye olarak belirlenmiştir.

Tablo 3.35.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modellerin XGBoost ve RF olduğu görülmektedir. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. 1D-CNN modeli 0,833 MCC değeri ile yüksek doğrulukta tahminleme yapmaktadır. KNN modeli başarı kriteri olan 0,75 MCC değerini geçerken MLP modeli ise başarısız olmuştur. Sonuç olarak, 2 saniyelik zaman ölçeğinde Gaint Analysis testi esnasında toplanan ivmeölçer verileri ve EKG verileri birlikte analiz edilerek Kırılgnlık tahminlemesinde XGBoost, RF ve 1D-CNN modeli başarılı olmuştur.

Tablo 3.36.'da verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin birbirine yakın sonuçlarla RF ve XGBoost olduğu saptanmıştır. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. 1D-CNN modeli başta olmak üzere MLP ve KNN modelleri de yüksek oranda doğru tahminlemeler yaparak başarılı olmuştur. Sonuç olarak, 4 saniyelik zaman ölçeğinde Gaint Analysis testi esnasında toplanan ivmeölçer verileri ve EKG verileri birlikte analiz edilerek Kırılgnlık tahminlemesinde Naive Bayes modeli haricindeki tüm modeller başarılı olmuştur.

Tablo 3.37.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelin birbirine yakın sonuçlarla RF ve XGBoost olduğu görülmüştür. Bu modeller kesin sonuca yakın doğrulukta tahminleme yapmaktadır. Sırasıyla 1D-CNN, KNN ve MLP modelleri e yüksek oranda doğru tahminlemeler yaparak başarılı olmuştur. Sonuç olarak, 8 saniyelik zaman ölçeğinde Gaint Analysis testi esnasında toplanan ivmeölçer verileri ve EKG verileri birlikte analiz edilerek Kırılgnlık tahminlemesinde Naive Bayes modeli haricindeki tüm modeller başarılı olmuştur.

Tablo 3.35., 3.36. ve 3.37. sonuçları incelendiğinde, 2 saniyelik zaman ölçeği 4 saniyeye çıkarıldığında KNN, MLP ve 1D-CNN modellerinin başarı metriklerinde kayda değer bir yükseliş gözlemlenmiştir. XGBoost ve RF modelinde büyük bir iyileşme görülmemesinin nedeni zaten bu modellerin başarı metriklerinin 1'e yakınsıyor olmasındandır. 4 saniyelik ölçeğin 8 saniyeye çıkarılmasıyla başarı metriklerinde yaşanan

yükselişler 2 saniye ölçeğinden 4 saniyelik ölçeğe çıkıldığındaki kadar anlamlı değildir. Sonuç olarak, Gaint Analysis testi esnasında toplanan ivmeölçer verileri ve EKG verileri birlikte analiz edilerek Kırılgnlık tahminlemesinde en uygun zaman ölçeği 4 saniye olarak belirlenmiştir.

Gaint Analysis testi genel olarak modellerin başarılı tahminleme yapmada zorlandığı testlerden biri olmuştur. 2 saniye zaman ölçeğinde ivmeölçer verilerinin analizinde (Tablo 3.30) XGBoost modelinin MCC değeri 0,80 bandının altına kadar düşmüştür. En başarılı sonucu veren veri seti ivmeölçer verilerinin EKG verileriyle birleştirilmesiyle oluşan veri setleri olmuştur. İkinci sırada ise EKG verilerinden oluşan veri seti gelmektedir.

#### e) 6MWT

Tablo 3.38.'de verilen performans başarı metrikleri sonuçlarına göre, en iyi sonucu veren modelin RF olduğu görülmektedir. RF modeli ile en iyi ikinci model olan XGBoost modeli yüksek doğruluk oranında tahminleme yapmaktadır. KNN, MLP ve 1D-CNN modelleri ise kırılgnlık tahminlemesinde başarısız olmuşlardır. Sonuç olarak, 2 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan ivmeölçer verileri analiz edilerek Kırılgnlık tahminlemesinde sadece RF ve XGBoost modelleri başarılı olabilmıştır.

Tablo 3.39.'da verilen performans başarı metrikleri sonuçlarına göre, en iyi sonucu veren modelin RF olduğu görülmektedir. RF modeli ile en iyi ikinci model olan XGBoost modeli yüksek doğruluk oranında tahminleme yapmaktadır. Diğer modeller KNN, MLP ve 1D-CNN ise kırılgnlığı doğru bir şekilde tahmin edememektedir. Sonuç olarak, 4 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan ivmeölçer verileri analiz edilerek Kırılgnlık tahminlemesinde sadece RF ve XGBoost modelleri başarılı olabilmıştır.

Tablo 3.40.'da verilen performans başarı metrikleri sonuçlarına göre, en iyi sonucu veren modelin RF olduğu görülmektedir. RF modeli ile en iyi ikinci model olan XGBoost modeli yüksek doğruluk oranında tahminleme yapmaktadır. Sırasıyla KNN ve 1D-CNN modelleri başarı kriteri olan 0,75 MCC değerine yaklaşmış olsa da bu değere ulaşamayarak başarısız olmuştur. Sonuç olarak, 8 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan ivmeölçer verileri analiz edilerek Kırılgnlık tahminlemesinde sadece RF ve XGBoost modelleri başarılı olabilmıştır.

Tablo 3.38., 3.39. ve 3.40. sonuçları incelendiğinde, her 3 zaman ölçeğinde de sadece RF ve XGBoost modellerinin başarılı olabildiği ölçek değıştikçe model başarı metriklerinde artış olduđu ancak bu artışın 0,75 MCC değerini aşmaya yetmediği gözlenmiştir. Sonuç olarak, 6MWT testi esnasında toplanan ivmeölçer verileri analiz edilerek Kırılgnlık tahminlemede ölçek değışimlerinin RF ve XGBoost üzerinde sınırlı etkisi olduđu da göz önünde bulundurularak 4 saniyelik zaman ölçeğinin bu problem için daha uygun olduđu kanısına varılmıştır. Bu kanıya varılma nedeni 2 saniyelik zaman ölçeğinden 4 saniyeye çıkıldığında RF ve XGBoost'ta minimal bir yükseliş görülmüş ancak ölçek 8 saniyeye çıkarıldığında RF modelinde yine minimalde olsa bir düşüş gözlenmiş olmasındandır.

Tablo 3.41.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller birbirlerine yakın MCC değerleriyle XGBoost ve RF modelleri olmuştur. Bu modeller 0,94 üzeri MCC değerleriyle kesine yakın doğrulukta tahminleme yapabilmektedir. XBoost ve RF modellerinde bu denli yüksek başarı metrikleri alınmış olmasına rağmen KNN, MLP ve 1D-CNN modelleri 0,75 MCC sınırının altında kalarak başarısız olmuştur. Sonuç olarak, 2 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan EKG verileri analiz edilerek Kırılgnlık tahminlemede sadece RF ve XGBoost modelleri başarılı olabilmıştır.

Tablo 3.42.'de verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller birbirlerine yakın MCC değerleriyle RF ve XGBoost modelleri olmuştur. Bu modeller 0,96 üzeri MCC değerleriyle kesine yakın doğrulukta tahminleme yapabilmesine karşın KNN, MLP ve 1D-CNN modelleri beklenen başarı kriterine ulaşamayarak başarısız olmuştur. Sonuç olarak, 4 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan EKG verileri analiz edilerek Kırılgnlık tahminlemede sadece RF ve XGBoost modelleri başarılı olabilmıştır.

Tablo 3.43.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modeller birbirlerine yakın MCC değerleriyle RF ve XGBoost modelleri olmuştur. Bu modeller 0,97 üzeri MCC değerleriyle kesine yakın doğrulukta tahminleme yapabilmesine karşın KNN ve MLP modelleri 0,7 MCC değerini geçmiş ancak beklenen başarı kriterine ulaşamayarak başarısız olmuştur. 1D-CNN modeli ise diğer modellerin oldukça gerisinde kalmıştır. Sonuç olarak, 8 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan EKG verileri analiz edilerek Kırılgnlık tahminlemede sadece RF ve XGBoost modelleri başarılı olabilmıştır.

Tablo 3.41., 3.42. ve 3.43. sonuçları incelendiğinde, kurulan modellerin bu test özelinde oldukça zorlandığı görülmektedir. Tüm zaman ölçeklerinde XGBoost ve RF hariç tamamının başarısız olduğu açık bir şekilde ortadadır. 2 saniyelik zaman ölçeğinden 8 saniyelik zaman ölçeğine çıkıldıkça tüm model performanslarında yükseliş gözlemlense de bu yükseliş modelleri başarılı olarak nitelendirmeye yetmemiştir. Bu nedenle zaman ölçeği anlamında anlamlı bir sonuç elde edilememiştir. 6MWT testi esnasında toplanan EKG verileri analiz edilerek Kırılgnlık tahminlemesinde zamana karar verebilmek amacıyla başarılı modellerin en yüksek MCC değerine ulaştığı ölçeğe bakılarak 8 saniyelik zaman ölçeğinde karar kılınmıştır.

Tablo 3.44.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modellerin RF ve XGBoost olduğu görülmektedir. Bu modeller 0,98 MCC değeriyle kesine yakın doğrulukta tahminleme yapmaktadır. 1D-CNN Modeli yüksek doğruluk oranıyla tahminleme yapabilirken, sırasıyla KNN ve MLP modelleri başarı kriterini aşarak başarılı olmuşlardır. Sonuç olarak, 2 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan üç eksenli, İvmeölçer verileri ile EKG verilerinin birlikte analiz edilerek Kırılgnlık tahminlemesinde Naive Bayes modeli haricinde tüm modeller başarılı olmuştur.

Tablo 3.45.'te verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelinin RF ve XGBoost olduğu görülmektedir. Bu modeller 0,99 MCC değeriyle kesine yakın doğrulukta tahminleme yapabilmektedir. Sırasıyla 1D-CNN, KNN ve MLP ise yüksek doğruluk oranlarıyla tahminleme yapabildiği görülmektedir. Sonuç olarak, 4 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan üç eksenli, İvmeölçer verileri ile EKG verilerinin birlikte analiz edilerek Kırılgnlık tahminlemesinde Naive Bayes modeli haricinde tüm modeller başarılı olmuştur.

Tablo 3.46.'da verilen performans başarı metrikleri sonuçlarına göre, en başarılı modelinin RF ve XGBoost olduğu görülmektedir. Bu modeller 0,99 MCC değeriyle kesine yakın doğrulukta tahminleme yapabilmektedir. Sırasıyla 1D-CNN, KNN ve MLP ise yüksek doğruluk oranlarıyla tahminleme yapabildiği görülmektedir. Sonuç olarak, 8 saniyelik zaman ölçeğinde 6MWT testi esnasında toplanan üç eksenli, İvmeölçer verileri ile EKG verilerinin birlikte analiz edilerek Kırılgnlık tahminlemesinde Naive Bayes modeli haricinde tüm modeller başarılı olmuştur.

Tablo 3.44., 3.45. ve 3.46. sonuçları incelendiğinde, 2 saniyelik zaman ölçeği 4 saniyeye çıkartıldığında tüm modeller için kayda değer bir artış olduğu özellikle KNN ve MLP modellerinin yüksek doğrulukla tahminleme yapabilme kabiliyeti kazandığı görülmektedir. 4 saniyelik ölçek 8 saniyeye çıkarıldığında ise RF ve XGBoost modelleri kayda değer bir gelişme gösteremezken KNN ve 1D-CNN modellerinin performansı artmış ancak MLP performansında düşüş gözlenmiştir. Sonuç olarak, 6MWT testi esnasında toplanan üç eksenli, İvmeölçer verileri ile EKG verilerinin birlikte analiz edilerek Kırılgnalık tahminlemede en verimli zaman ölçeğinin 4 saniyelik zaman ölçeği olduğu değerlendirilmiştir.

6MWT testi genel olarak ML modellerinin en çok zorlandığı test olmuştur. Gerek ivmeölçer verileri gerekse EKG verileri yalnız başlarına kullanıldığında XGBoost ve RF modelleri haricinde başarıya ulaşan bir model olmamıştır. Özellikle ivmeölçer verilerinin yalnız kullanıldığı senaryolarda KNN, MLP ve 1D-CNN modelleri başarı kriteri olan 0,75 MCC değerinin oldukça altında kaldığı saptanmıştır. Bu durum ancak ivmeölçer ve EKG verilerinin birlikte kullanıldığı senaryoda düzelmiş ve tüm modellerin başarılı olduğu sonuçlar vermiştir. Sonuç olarak, 6MWT testinde en iyi tahminlemenin yapıldığı senaryo 4 saniyelik zaman ölçeğinde ivmeölçer ve EKG verilerinin birlikte analiz edilerek kırılgnalık tahmini yaptığı senaryo olmuştur.

Yapılan incelemeler sonucunda oluşturulan Tablo 4.1. Özet Sonuçlar'da bir araya getirilmiştir. Stair, TUG, Veloergometri gibi testlerde 2 saniyelik zaman ölçeği seçilirken, Gaint Analysis ve 6MWT gibi daha uzun soluklu testlerde 4 saniyelik zaman ölçeğinin başarılı olduğu görülmüştür. Veri tipinde ivme ölçer ve EKG verilerinin birlikte kullanımının tüm fiziksel testler esnasında başarıyı arttırdığı ve yine tün fiziksel aktiviteler esnasında en başarılı modellerin XGBoost ve RF olduğu saptanmıştır.

**Tablo 4.1.** Özet Sonuçlar

| Fiziksel Aktivite Testi | En Uygun Zaman Ölçeği | En Uygun Veri Tipi | En İyi Sonucu Veren Model |
|-------------------------|-----------------------|--------------------|---------------------------|
| Stair                   | 2 Saniye              | İvmeölçer & EKG    | XGBoost, RF               |
| TUG                     | 2 Saniye              | İvmeölçer & EKG    | XGBoost, RF               |
| Veloergometri           | 2 Saniye              | İvmeölçer & EKG    | XGBoost, RF               |
| Gaint Analysis          | 4 Saniye              | İvmeölçer & EKG    | XGBoost, RF               |
| 6MWT                    | 4 Saniye              | İvmeölçer & EKG    | XGBoost, RF               |

Bu tez kapsamında yapılan çalışma genel olarak incelendiğinde, oldukça ilginç ve dikkat çekici sonuçların elde edildiği görülmektedir. Kırılğanlığın fiziksel ilişkisi birçok çalışmada ele alınmış ve bu nedenle çalışmalarda ağırlıklı olarak İvmeölçer verileri tercih edilmiştir. İvmeölçer, bireyin hareketlerini ve bu hareketler esnasındaki akıcılığı ölçerek, bu akıcılığın kırılğanlık ile ilişkisini tahmin etmeyi hedeflemektedir. Ancak, bu yaklaşımda bireyin hareketi esnasında kalbin verdiği tepkiler göz ardı edilmektedir.

Bu çalışmanın en önemli bulgusu, fiziksel aktivite testleri sırasında toplanan EKG verilerinin kırılğanlığın tahmin edilmesinde daha verimli sonuçlar verdiğini ortaya koymaktadır. Bununla birlikte, daha kesin sonuçlara ulaşabilmek adına, İvmeölçer ve EKG verilerinin birlikte kullanılması gerektiği vurgulanmaktadır. Çalışmada kullanılan ML modelleri arasında, RF ve XGBoost modelleri her senaryoda açık ara ile en yüksek başarıyı sağlamıştır.

Zaman ölçeği açısından bakıldığında, Stair, TUG ve Veloergometri testleri için en uygun zaman ölçeği 2 saniye olarak belirlenmiştir. Bu durum, bu tür aktivitelerde hızlı ve ani dinamik değişimlerin meydana geldiğini, dolayısıyla daha kısa zaman dilimlerinin bu değişimleri yakalamada yeterli olduğunu göstermektedir. Öte yandan, Gait Analysis ve 6MWT gibi daha uzun süreli ya da daha yavaş dinamiklere sahip aktivitelerde ise 4 saniyelik zaman ölçeklerinin daha uygun olduğu görülmüştür. Bu bulgular, farklı fiziksel aktivitelerin doğasına uygun zaman aralıklarının belirlenmesinin model performansını optimize etmek açısından kritik bir rol oynadığını ortaya koymaktadır. Kısa süreli ve hızlı değişimlere sahip aktiviteler için daha dar bir zaman ölçeği tercih edilirken, daha uzun süreli ve yavaş dinamiklere sahip aktivitelerde daha geniş zaman ölçeklerinin kullanılması, elde edilen verilerin anlamlılığını ve modelin tahmin başarısını artırmaktadır. Bu yaklaşım, kırılğanlık tahmini çalışmalarında zaman ölçeği seçiminin önemini vurgulamakta ve farklı fiziksel aktiviteler için özelleştirilmiş parametrelerin kullanılmasının gerekliliğini işaret etmektedir.

## 5. SONUÇ

Bu tez kapsamında yapılan çalışma, yaşlı bireylerde kırılgnlık tespitine yönelik çeşitli ML algoritmalarının performansları değerlendirilmiş ve kırılgnlık tahmini için biyofiziksel verilerin önemini vurgulanmıştır. Çalışmada elde edilen bulgular, kırılgnlığın fiziksel aktivitelerle olan ilişkisini inceleyen önceki çalışmalardan farklı olarak, EKG verilerinin kırılgnlık tahmininde daha verimli sonuçlar sağladığını ortaya koymuştur.

Literatürde yaygın olarak ivmeölçer verileri kullanılarak kırılgnlık tahmin edilmekteydi, çünkü ivmeölçerler bireyin hareketlerini ve hareket akıcılığını ölçerek kırılgnlığı tespit etmeye çalışmaktadır. Ancak bu çalışmada, fiziksel aktivite sırasında kalbin verdiği tepkilerin (EKG sinyalleri) göz ardı edilmemesi gerektiği anlaşılmıştır. Özellikle fiziksel aktivite testleri sırasında toplanan EKG verilerinin, kırılgnlık tahmininde yalnızca ivmeölçer verilerine kıyasla daha güçlü sonuçlar verdiği görülmüştür. Bu bulgu, kırılgnlığın daha doğru bir şekilde tespit edilmesi için ivmeölçer ve EKG verilerinin birlikte kullanılması gerektiğini işaret etmektedir.

Çalışmada kullanılan RF ve XGBoost algoritmaları, her iki veri türünde de tutarlı bir şekilde en yüksek başarıyı sağlamış ve kırılgnlık tespitinde güvenilir modeller olarak ön plana çıkmıştır. Ancak, zaman ölçeğinin model performansı üzerindeki etkileri de dikkatle ele alınmalıdır. Yapılan analizlerde, 2 saniyelik zaman ölçeği özellikle Stair, TUG ve Veloergometri gibi hızlı ve ani dinamiklerin yoğun olduğu aktiviteler için daha uygun bulunurken, Gait Analysis ve 6MWT gibi daha uzun süreli ve daha yavaş dinamiklere sahip aktivitelerde ise 4 saniyelik zaman ölçeği tercih edilmiştir. Bu bulgular, farklı aktivitelerin doğasına uygun zaman ölçeklerinin belirlenmesinin model performansını optimize etmede kritik bir rol oynadığını ortaya koymaktadır. Bu bağlamda, zaman ölçeği, kırılgnlık tahmini çalışmalarında model başarısını artırmak için dikkatle ele alınması gereken önemli bir parametre olarak değerlendirilmelidir. Her bir aktivite türüne özgü en uygun zaman ölçeğinin belirlenmesi, model performansını iyileştirmek ve genel başarıyı maksimize etmek açısından büyük önem taşımaktadır.

Sonuç olarak, bu çalışma, yaşlı bireylerde kırılgnlık tespitinde EKG verilerinin önemini ortaya koymuş ve ML algoritmalarının kırılgnlık tahminindeki başarısını göstermiştir. İvmeölçer ve EKG verilerinin birlikte kullanılması, daha güvenilir sonuçlar elde edilmesini sağlamaktadır. Gelecek çalışmalar, bu iki veri tipinin entegrasyonu üzerine

yoğunlaşarak, kırılma tespiti için daha kapsamlı ve etkili modellerin geliştirilmesine katkıda bulunabilir.



## 6. KAYNAKLAR

- [1] E. Hoogendijk, J. Afilalo, K. Ensrud, P. Kowal, G. Onder and L. Fried, "Frailty: implications for clinical practice and public health", *The Lancet*, vol. 394, no. 10206, pp. 1365-1375, 2019. Available: 10.1016/s0140- 6736(19)31786-6.
- [2] Y. N. Panhwar, F. Naghdy, G. Naghdy, D. Stirling, ve J. Potter, "Assessment of frailty: a survey of quantitative and clinical methods," *BMC Biomedical Engineering*, vol. 1, p. 7, 2019. DOI: 10.1186/s42490-019-0007-y.
- [3] D. B. Rolfson, S. R. Majumdar, R. T. Tsuyuki, A. Tahir, ve K. Rockwood, "Validity and reliability of the Edmonton Frail Scale," *Age and Ageing*, vol. 35, no. 5, pp. 526-529, 2006. DOI: 10.1093/ageing/afl041.
- [4] Fried, L. P., C. M. Tangen, J. Walston, A. B. Newman, C. Hirsch, J. Gottdiener, T. Seeman, R. Tracy, W. J. Kop, G. Burke, ve M. A. McBurnie, "Frailty in older adults: Evidence for a phenotype," *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, vol. 56, no. 3, pp. M146-M156, 2001. DOI: 10.1093/gerona/56.3.M146.
- [5] A. Martínez-Ramírez, P. Lecumberri, M. Gómez, L. Rodríguez-Mañas, F. J. García, ve M. Izquierdo, "Frailty assessment based on wavelet analysis during quiet standing balance test," *Journal of Biomechanics*, vol. 44, no. 12, pp. 2213–2220, 2011. DOI: 10.1016/j.jbiomech.2011.06.007.
- [6] N. Millor, P. Lecumberri, M. Gómez, A. Martínez-Ramírez, ve M. Izquierdo, "An evaluation of the 30-s chair stand test in older adults: frailty detection based on kinematic parameters from a single inertial unit," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, p. 86, 2013. DOI: 10.1186/1743-0003-10-86.
- [7] A. Galán-Mercant ve A. I. Cuesta-Vargas, "Differences in trunk kinematic between frail and nonfrail elderly persons during turn transition based on a smartphone inertial sensor," *BioMed Research International*, vol. 2013, p. 279197, 2013. DOI: 10.1155/2013/279197.
- [8] A. Galán-Mercant ve A. I. Cuesta-Vargas, "Differences in trunk accelerometry between frail and nonfrail elderly persons in sit-to-stand and stand-to-sit transitions based on a mobile

inertial sensor," *JMIR Mhealth Uhealth*, vol. 1, no. 2, p. e21, 2013. DOI: 10.2196/mhealth.2710.

[9] B. R. Greene, E. P. Doheny, R. A. Kenny, ve B. Caulfield, "Classification of frailty and falls history using a combination of sensor-based mobility assessments," *Physiological Measurement*, vol. 35, no. 10, pp. 2053–2066, 2014. DOI: 10.1088/0967-3334/35/10/2053.

[10] B. R. Greene, E. P. Doheny, A. O'Halloran, ve R. A. Kenny, "Frailty status can be accurately assessed using inertial sensors and the TUG test," *Age and Ageing*, vol. 43, no. 3, pp. 406–411, 2014. DOI: 10.1093/ageing/aft176.

[11] F. M. Jansen, R. G. Prins, A. Etman, H. P. van der Ploeg, S. I. de Vries, F. J. van Lenthe, ve F. H. Pierik, "Physical activity in non-frail and frail older adults," *PLoS ONE*, vol. 10, no. 4, p. e0123168, 2015. DOI: 10.1371/journal.pone.0123168.

[12] S. Chen, T. Honda, T. Chen, K. Narazaki, Y. Haeuchi, A. Supartini, ve S. Kumagai, "Screening for frailty phenotype with objectively-measured physical activity in a west Japanese suburban community: evidence from the Sasaguri Genkimon Study," *BMC Geriatrics*, vol. 15, no. 36, 2015. DOI: 10.1186/s12877-015-0037-9.

[13] D. Castaneda-Gameros, S. Redwood, ve J. L. Thompson, "Physical activity, sedentary time, and frailty in older migrant women from ethnically diverse backgrounds: a mixed-methods study," *Journal of Aging and Physical Activity*, 2017. DOI: 10.1123/japa.2016-0287.

[14] M. Huisingh-Scheetz, K. Wroblewski, M. Kocherginsky, E. Huang, W. Dale, L. Waite, ve L. P. Schumm, "The relationship between physical activity and frailty among U.S. older adults based on hourly accelerometry data," *Journals of Gerontology: Medical Sciences*, vol. 73, no. 5, pp. 622–629, 2018. DOI: 10.1093/gerona/glx208.

[15] A. Yuki, R. Otsuka, C. Tange, Y. Nishita, M. Tomida, F. Ando, H. Shimokata, ve H. Arai, "Daily physical activity predicts frailty development among community-dwelling older Japanese adults," *Journal of the American Medical Directors Association*, vol. 20, no. 8, pp. 1032–1036.e2, 2019. DOI: 10.1016/j.jamda.2019.01.001.

- [16] C. P. Jansen, N. Toosizadeh, M. J. Mohler, B. Najafi, C. Wendel, ve M. Schwenk, "The association between motor capacity and mobility performance: frailty as a moderator," *European Review of Aging and Physical Activity*, vol. 16, no. 16, 2019. DOI: 10.1186/s11556-019-0223-4.
- [17] C. Ziller, T. Braun, ve C. Thiel, "Frailty phenotype prevalence in community-dwelling older adults according to physical activity assessment method," *Clinical Interventions in Aging*, vol. 15, pp. 343–355, 2020. DOI: 10.2147/CIA.S238204.
- [18] H. Kikuchi, S. Inoue, S. Amagasa, N. Fukushima, M. Machida, H. Murayama, T. Fujiwara, S. Chastin, N. Owen, ve Y. Shobugawa, "Associations of older adults' physical activity and bout-specific sedentary time with frailty status: compositional analyses from the NEIGE study," *Experimental Gerontology*, 2020. DOI: 10.1016/j.exger.2020.111149.
- [19] O. Theou, J. M. Jakobi, A. A. Vandervoort, ve G. R. Jones, "A comparison of physical activity (PA) assessment tools across levels of frailty," *Archives of Gerontology and Geriatrics*, vol. 54, no. 3, pp. e307–e314, 2012. DOI: 10.1016/j.archger.2011.12.005.
- [20] M. Schwenk, J. Mohler, C. Wendel, K. D'Huyvetter, M. Fain, R. Taylor-Piliae, ve B. Najafi, "Wearable sensor-based in-home assessment of gait, balance, and physical activity for discrimination of frailty status: baseline results of the Arizona Frailty Cohort Study," *Gerontology*, vol. 61, no. 3, pp. 258–267, 2015. DOI: 10.1159/000369095.
- [21] N. Toosizadeh, J. Mohler, ve B. Najafi, "Assessing upper extremity motion: an innovative method to identify frailty," *Journal of the American Geriatrics Society*, vol. 63, no. 6, pp. 1181–1186, 2015. DOI: 10.1111/jgs.13451.
- [22] S. Parvaneh, J. Mohler, N. Toosizadeh, G. S. Grewal, ve B. Najafi, "Postural transitions during activities of daily living could identify frailty status: application of wearable technology to identify frailty during unsupervised condition," *Gerontology*, vol. 63, no. 5, pp. 479–487, 2017. DOI: 10.1159/000460292.
- [23] J. Razjouyan, A. D. Naik, M. J. Horstman, M. E. Kunik, M. Amirmazaheri, H. Zhou, A. Sharafkhaneh, ve B. Najafi, "Wearable sensors and the assessment of frailty among vulnerable older adults: an observational cohort study," *Sensors*, vol. 18, no. 5, p. 1336, 2018. DOI: 10.3390/s18051336.

- [24] N. Straiton, M. Alharbi, A. Bauman, L. Neubeck, J. Gullick, R. Bhindi, ve R. Gallagher, "The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review," *Maturitas*, vol. 112, pp. 85–93, 2018. DOI: 10.1016/j.maturitas.2018.03.016.
- [25] H. Lee, B. Joseph, A. Enriquez, ve B. Najafi, "Toward using a smartwatch to monitor frailty in a hospital setting: using a single wrist-wearable sensor to assess frailty in bedbound inpatients," *Gerontology*, vol. 64, no. 4, pp. 389–400, 2018. DOI: 10.1159/000484241.
- [26] N. Toosizadeh, B. Joseph, M. R. Heusser, T. Orouji Jokar, J. Mohler, H. A. Phelan, ve B. Najafi, "Assessing upper-extremity motion: an innovative, objective method to identify frailty in older bed-bound trauma patients," *Journal of the American College of Surgeons*, 2016. DOI: 10.1016/j.jamcollsurg.2016.03.030.
- [27] A. P. Hassler, E. Menasalvas, F. J. García-García, L. Rodríguez-Mañas, ve A. Holzinger, "Importance of medical data preprocessing in predictive modeling and risk factor discovery for the frailty syndrome," *BMC Medical Informatics and Decision Making*, vol. 19, no. 33, 2019. DOI: 10.1186/s12911-019-0747-6.
- [28] M.-A. Le Pogam, L. Seematter-Bagnoud, T. Niemi, D. Assouline, N. Gross, B. Trachsel, ve B. Santos-Eggimann, "Development and validation of a knowledge-based score to predict Fried's frailty phenotype across multiple settings using one-year hospital discharge data: The electronic frailty score," *EClinicalMedicine*, vol. 44, p. 101260, 2022. DOI: 10.1016/j.eclinm.2021.101260.
- [29] D. E. C. Leme ve C. de Oliveira, "Machine learning models to predict future frailty in community-dwelling middle-aged and older adults: The ELSA cohort study," *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, vol. 78, no. 11, pp. 2176–2184, 2023. DOI: 10.1093/gerona/glad127.
- [30] R. Szczepanowski, I. Uchmanowicz, A. H. Pasieczna, J. Sobocki, R. Katarzyniak, G. Kołaczek, W. Lorkiewicz, M. Kędras, A. Dixit, J. Biegus, M. Wleklik, R. J. J. Gobbens, L. Hill, T. Jaarsma, A. Hussain, M. Barbagallo, ve F. C. Morabito, "Application of machine learning in predicting frailty syndrome in patients with heart failure," *Advances in Clinical and Experimental Medicine*, vol. 33, no. 3, pp. 309–315, 2024. DOI: 10.17219/acem/184040.

- [31] Y. Mizuguchi, M. Nakao, T. Nagai, ve diğerleri, "Machine learning–based gait analysis to predict clinical frailty scale in elderly patients with heart failure," *European Heart Journal - Digital Health*, vol. 5, no. 2, pp. 152–162, 2024. DOI: 10.1093/ehjdh/ztad082.
- [32] Y. Chen, Y. Yu, D. Yang, W. Zhang, V. Kouritas, ve X. Chen, "Developing and validating machine learning-based prediction models for frailty occurrence in those with chronic obstructive pulmonary disease," *Journal of Thoracic Disease*, vol. 16, no. 4, pp. 2482–2498, 2024. DOI: 10.21037/jtd-24-416.
- [33] Ç. B. Erdaş ve D. Ölçer, "Prediction of frailty grade using machine learning models," *2022 Medical Technologies Congress (TIPTEKNO)*, Antalya, Turkey, 2022, ss. 1-4, doi: 10.1109/TIPTEKNO56568.2022.9960172.
- [34] D. Sokas, M. Butkuvienė, E. Tamulevičiūtė-Prascienė, A. Beigienė, R. Kubilius, A. Petrėnas, ve B. Paliakaitė, "Wearable-based signals during physical exercises from patients with frailty after open-heart surgery (version 1.0.0)," *PhysioNet*, 2022. DOI: 10.13026/mp8k-7p27.
- [35] D. Figo, P. C. Diniz, D. R. Ferreira, ve J. M. P. Cardoso, "Preprocessing techniques for context recognition from accelerometer data," *Personal and Ubiquitous Computing*, vol. 14, no. 7, pp. 645–662, 2010. DOI: 10.1007/s00779-010-0293-9.
- [36] S. Barrash, Y. Shen, and G. B. Giannakis, "Scalable and adaptive KNN for regression over graphs," in 2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Dec. 2019, pp. 241–245, DOI: 10.1109/CAMSAP45676.2019.9022509.
- [37] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324.
- [38] S. S. Nain, D. Garg, and S. Kumar, "Performance evaluation of the WEDM process of aeronautics super alloy," *Materials and Manufacturing Processes*, vol. 33, no. 16, pp. 1793–1808, 2018, doi: 10.1080/10426914.2018.1476761

- [39] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 2016, pp. 785–794. DOI: 10.1145/2939672.2939785.
- [40] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2015, DOI: 10.1109/TBME.2015.2468589
- [41] K. Singh, A. Mahajan, ve V. Mansotra, "1D-CNN based model for classification and analysis of network attacks," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, pp. 604–613, 2021. DOI: 10.14569/IJACSA.2021.0121169.
- [42] A. V. D. Sano, A. A. Stefanus, E. D. Madyatmadja, H. Nindito, A. Purnomo, ve C. P. M. Sianipar, "Proposing a visualized comparative review analysis model on tourism domain using Naïve Bayes classifier," *Procedia Computer Science*, vol. 227, pp. 482–489, 2023. DOI: 10.1016/j.procs.2023.10.549.
- [43] P. C. Cheeseman and J. C. Stutz, "Bayesian classification (AutoClass): theory and results," in *Advances in Knowledge Discovery and Data Mining*, vol. 180, pp. 153–180, 1996.
- [44] Y. Wang, S. Wei, R. Zuo, ve diğerleri, "Automatic and real-time tissue sensing for autonomous intestinal anastomosis using hybrid MLP-DC-CNN classifier-based optical coherence tomography," *Biomedical Optics Express*, vol. 15, no. 4, pp. 2543–2557, 2024. DOI: 10.1364/BOE.521652.
- [45] S. Prusty, S. Patnaik, ve S. K. Dash, "SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer," *Frontiers in Nanotechnology*, vol. 4, p. 972421, 2022. DOI: 10.3389/fnano.2022.972421.
- [46] Ö. Aydemir, "Dengesiz veri kümelerinin sınıflandırılmasında poligon alan metriğinin sınıflandırıcı performans değerlendirilmesi için kullanılması," *Yüüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, vol. 27, no. 2, pp. 194–205, 2022. DOI: 10.53433/yyufbed.1066340.

[47] Ç. B. Erdaş, E. Sümer, ve S. Kibarođlu, "Neurodegenerative disease detection and severity prediction using deep learning approaches," *Biomedical Signal Processing and Control*, vol. 70, p. 103069, 2021. DOI: 10.1016/j.bspc.2021.103069.

[48] A. Alan ve M. Karabatak, "Veri seti - sınıflandırma ilişkisinde performansa etki eden faktörlerin deđerlendirilmesi," *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 32, no. 2, pp. 531–540, 2020. DOI: 10.35234/fumbd.738007.

