



**SPOR BİLİMLERİNDE KULLANILAN Y-DENGE  
VERİLERİNİN MAKİNE ÖĞRENİMİ YÖNTEMLERİ  
İLE ANALİZİ**

**Süheda AKDAĞ**

**2023  
YÜKSEK LİSANS TEZİ  
BİLGİSAYAR MÜHENDİSLİĞİ**

**Tez Danışmanı  
Doç. Dr. Caner ÖZCAN**

**SPOR BİLİMLERİNDE KULLANILAN Y-DENGE VERİLERİNİN  
MAKİNE ÖĞRENİMİ YÖNTEMLERİ İLE ANALİZİ**

**Süheda AKDAĞ**

**T.C.  
Karabük Üniversitesi  
Lisansüstü Eğitim Enstitüsü  
Bilgisayar Mühendisliği Anabilim Dalında  
Yüksek Lisans Tezi  
Olarak Hazırlanmıştır**

**Tez Danışmanı  
Doç. Dr. Caner ÖZCAN**

**KARABÜK  
Ağustos 2023**

Süheda AKDAĞ tarafından hazırlanan “SPOR BİLİMLERİNDE KULLANILAN Y-DENGE VERİLERİNİN MAKİNE ÖĞRENİMİ YÖNTEMLERİ İLE ANALİZİ ” başlıklı bu tezin Yüksek Lisans Tezi olarak uygun olduğunu onaylarım.

Doç. Dr. Caner ÖZCAN .....

Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışma, jürimiz tarafından Oy Birliği ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir. 28/08/2023

Ünvanı, Adı SOYADI (Kurumu)

İmzası

Başkan : Dr. Öğr. Üyesi Emel SOYLU (SAMÜ)

ONLINE

Üye : Doç. Dr. Caner ÖZCAN (KBÜ)

.....

Üye : Dr. Öğr. Üyesi Mehmet Zahid YILDIRIM (KBÜ)

.....

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Yüksek Lisans derecesini onamıştır.

Doç. Dr. Zeynep ÖZCAN .....

Lisansüstü Eğitim Enstitüsü Müdürü

*“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”*

Süheda AKDAĞ

## **ÖZET**

**Yüksek Lisans Tezi**

### **SPOR BİLİMLERİNDE KULLANILAN Y-DENGE VERİLERİNİN MAKİNE ÖĞRENİMİ YÖNTEMLERİ İLE ANALİZİ**

**Süheda AKDAĞ**

**Karabük Üniversitesi**

**Lisansüstü Eğitim Enstitüsü**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Tez Danışmanı:**

**Doç. Dr. Caner ÖZCAN**

**Ağustos 2023, 61 sayfa**

Bu çalışmada, sporcuların dinamik denge ve fonksiyonel hareket yeteneklerini değerlendirmek için yaygın olarak kullanılan Y-Denge Testi (YDT) verilerinin makine öğrenimi yöntemleri ile analizi sağlanmıştır. YDT, sporcuların vücutlarını kontrol edebilme, denge sağlayabilme ve potansiyel sakatlık risklerini belirleyebilme yeteneklerini değerlendirerek, özelleştirilmiş antrenman programlarının tasarlanmasına ve performansın artırılmasına katkı sağlamasına yardımcı olan önemli bir yöntemdir. Testten doğru sonuçlar elde etmek için testi doğru şekilde yapmak çok önemlidir. YDT verilerinin hassas ölçümü her birey için ayrı ölçülmeli, test gridi ve ekipmanları doğru bir şekilde kurulmalıdır. Sporcuların performansını artırmak ve yaralanma risklerini azaltmak için eğitim programlarının tasarlanmasında yanlış bilgilerin kullanılmasına yol açabilecek yanıltıcı sonuçlardan kaçınmak için net talimatlar ve standart bir protokol izlenmelidir. Tüm bu süreç, testi yapan kişi

tarafından dikkatle ve özenle yapılmalıdır. Çalışma, YDT verilerini daha iyi anlamak, doğru içgörüler elde etmek ve farklı makine öğrenimi modellerinin YDT sonuçlarını tahmin etme yeteneğini değerlendirmek için bir çerçeve sunmaktadır. Bu kapsamlı yaklaşım, test yöneticilerine ölçümler ve sonuçlardaki olası tutarsızlıkları güvenle yönlendirmeleri için bir yol sunar. Çalışma kapsamında kullanılan YDT verileri ile ilgili yasal izinler alınmıştır ve çalışmada kullanılan veri kümesi, Uşak ili kapsamındaki farklı branşlarda faaliyet gösteren spor kulüplerinden elde edilmiştir. Çalışmada, bireysel ölçüm gerektiren YDT değerleri yaş, cinsiyet, antrenman yılı gibi özelliklere dayalı olarak farklı makine öğrenme yöntemleri ile ortaya çıkan tahmin sonuçları analiz edilmiştir. Sonuçlar, Ortalama Kare Hatası (MSE), R-Kare (R<sup>2</sup>) ve Kök Ortalama Kare Hatası (RMSE) ile değerlendirilmiştir. Her YDT verisi Decision Tree, Random Forest, AdaBoost, Gradient Boosting, XGBoost ve SGD algoritmaları ile ayrı bir şekilde test edilmiştir. Bu ayrı modellerin analizi sonucunda, en yüksek performans değerleri; MSE değeri 0,01, R<sup>2</sup> değeri 0,77 ve RMSE değeri 0,09 ile Gradient Boosting algoritması tarafından elde edilmiştir. Sonuçlar, tahmin edilen YDT değerlerinin, sporcuların performansını artırabilecek ve yaralanma risklerini azaltabilecek eğitim programlarının tasarlanmasına yardımcı olabileceğini göstermiştir. Genel olarak, çalışmanın bulguları, doğru YDT veri ölçümünün önemini ve belirli özelliklere sahip bir bireye dayalı olarak YDT değerlerini tahmin etmede makine öğrenimi yöntemlerinin potansiyelini vurgulamaktadır. Kullanılan parametrelere göre değişkenlik gösteren analizler sonucunda en iyi sonuçları XGBoost, Gradient Boosting ve Random Forest algoritmaları vermiştir. Bu yaklaşım, yaralanma risklerini en aza indirirken sporcuların denge ve hareketliliğini geliştirebilecek özel eğitim programları oluşturmak için koçlara, eğitmenlere ve sağlık uzmanlarına değerli bilgiler sağlayabilecek ve test sürecini hızlandırabilecektir.

**Anahtar Sözcükler :** Makine öğrenimi, denetimli öğrenme, regresyon, tahmin, analiz, spor bilimi, y-denge testi, denetimli öğrenme, karar ağaçları, rastgele orman, toplu öğrenme.

**Bilim Kodu :** 92431

## **ABSTRACT**

**M. Sc. Thesis**

### **ANALYSIS OF Y-BALANCE DATA USED IN SPORTS SCIENCES WITH MACHINE LEARNING ALGORITHMS**

**Süheda AKDAĞ**

**Karabük University  
Institute of Graduate Programs  
Department of Computer Engineering**

**Thesis Advisor:**

**Assoc. Prof. Dr. Caner ÖZCAN**

**August 2023, 61 pages**

In this study, the analysis of Y-Balance Test (YBT) data, commonly used to assess athletes' dynamic balance and functional movement abilities, has been conducted using machine learning techniques. YBT evaluates athletes' capacity to control their bodies, maintain balance, and identify potential injury risks, thereby aiding the design of personalized training programs and performance enhancement. Accurate administration of the test is crucial for obtaining reliable outcomes. Precise measurement of YBT data should be individualized, with proper setup of the testing grid and equipment. To prevent misleading results that could lead to incorrect information in designing training programs to improve athlete performance and reduce injury risks, clear instructions and a standardized protocol must be followed. This entire process demands careful and meticulous execution by the tester. The study provides a framework for a comprehensive approach to better understand YBT data,

gain accurate insights, and assess the predictive abilities of various machine learning models on YBT outcomes. This comprehensive approach offers a path for test administrators to confidently address potential inconsistencies in measurements and results. Legal permissions were obtained for the YBT data used in the study, and the dataset, derived from different sports clubs in Uşak province, was employed. Individual YBT values requiring personal measurements were analyzed using different machine learning methods based on characteristics such as age, gender, and years of training. Results were evaluated using Mean Squared Error (MSE), R-Squared (R<sup>2</sup>), and Root Mean Squared Error (RMSE). Each YBT dataset was individually tested with Decision Tree, Random Forest, AdaBoost, Gradient Boosting, XGBoost, and SGD algorithms. Through separate model analyses, the superior performance metrics MSE of 0.01, R<sup>2</sup> of 0.77, and RMSE of 0.09 were achieved by the Gradient Boosting algorithm. The results indicate that predicted YBT values can assist in designing training programs to enhance athletes' performance and mitigate injury risks. Overall, the study's findings underscore the significance of accurate YBT data measurement and highlight the potential of machine learning methods to predict YBT values based on specific individual attributes. Analyses varying with the used parameters resulted in XGBoost, Gradient Boosting, and Random Forest algorithms yielding the best outcomes. This approach can provide valuable insights to coaches, trainers, and health experts in crafting specialized training programs that improve balance and mobility while minimizing injury risks, potentially expediting the testing process.

**Key Word** : Machine learning, supervised learning, regression, prediction, analysis, sport science, y-balance test, decision trees, random forest, ensemble learning.

**Science Code** : 92431



## TEŐEKKÜR

Bu tez alıőmasının planlanmasında, araőtırılmasında, yürütülmesinde ve oluşumunda ilgi ve desteęini esirgemeyen, alıőma azminden ilham aldıęım, yönlendirme ve bilgilendirmeleriyle alıőmamı bilimsel temeller ışığında őekillendiren sayın hocam Do. Dr. Caner ÖZCAN'a sonsuz teőekkürlerimi sunarım.

alıőma süresince her zaman yanımda olan ve yoğun alıőmalarım sırasında beni neőelendiren, ilgi ve desteęini hiç esirgemeyen sevgili eőim Erkut AKDAĖ'a teőekkür ederim.

Bana her anımda destek olan ve alıőmaya teővik eden sevgili arkadaşlarım Dilara ÖZDEMİR, Elif MEŐECİ ve Buse Yaren KAZANGİRLER'e teőekkür ederim.

alıőmanın geliştirilmesinde ve veri kümesinin oluşturulmasında desteęini esirgemeyen sayın hocam Do. Dr. Bahar ATEŐ'e teőekkürlerimi sunarım.

Ek olarak bu tez alıőmamızı "KBUBAP-23-YL-040" proje numarası ile desteklemeye layık gören Karabük Üniversitesi Bilimsel Araőtırma Projeleri Koordinasyon Birimi'ne teőekkürlerimi sunarım.

Sevgili aileme manevi hiçbir yardımı esirgemedен yanımda oldukları için tüm kalbimle teőekkür ederim.

## İÇİNDEKİLER

	<b><u>Sayfa</u></b>
KABUL.....	ii
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER .....	ix
ŞEKİLLER DİZİNİ.....	xi
ÇİZELGELER DİZİNİ .....	xii
SİMGELER VE KISALTMALAR DİZİNİ .....	xiii
BÖLÜM 1 .....	1
GİRİŞ .....	1
BÖLÜM 2 .....	8
DENGE VE Y-DENGE TESTİ .....	8
2.1. DENGE .....	9
2.2. Y-DENGE TESTİ .....	10
2.2.1 Y-Denge Testi Uygulama Alanları .....	12
2.1.2. Y-Denge Testinin Uygulanması ve Ölçülmesi .....	13
2.3. Y-DENGE TESTİ VERİ KÜMESİ.....	15
BÖLÜM 3 .....	20
MAKİNE ÖĞRENİMİ VE REGRESYON .....	20
3.1. MAKİNE ÖĞRENİMİ .....	21
3.2. MAKİNE ÖĞRENİMİ REGRESYON ALGORİTMALARI.....	23
3.2.1. Karar Ağaçları (Decision Trees) Regresyon.....	25
3.2.2 Rastgele Orman (Random Forest) Regresyon .....	26
3.2.3. Adaptif Artırma (Adaptive Boosting) Regresyon.....	27

	<b><u>Sayfa</u></b>
3.2.4. Gradyan Artırma (Gradient Boosting) Regresyon.....	28
3.2.5. Aşırı Gradyan Artırma (Extreme Gradient Boosting) Regresyon .....	29
3.2.6. Stokastik Gradyan İniş (Stochastic Gradient Descent) Regresyon.....	30
BÖLÜM 4 .....	31
Y-DENGE TESTİ VERİLERİNİN MAKİNE ÖĞRENİMİ ALGORİTMALARI İLE ANALİZİ.....	31
4.1. VERİ ÖN İŞLEME .....	31
4.1.1. Veri Setini Yükleme ve Eksik Verilerin İşlenmesi.....	32
4.1.2 Normalizasyon İşlemi.....	34
4.1.3. Kategorik Verilerin İşlenmesi.....	36
4.2. MODEL SEÇİMİ .....	37
BÖLÜM 5 .....	38
DENEYSEL ÇALIŞMALAR .....	38
BÖLÜM 6 .....	52
SONUÇLAR VE ÖNERİLER .....	52
KAYNAKLAR .....	54
ÖZGEÇMİŞ .....	61

## ŞEKİLLER DİZİNİ

### Sayfa

Şekil 2.1. Y-denge testi kitinin kurulumu. ....	11
Şekil 2.2. Eller kalçada YDT (A) anterior uzanma (B) posterolateral uzanma (C) posteromedial uzanma [52]. ....	14
Şekil 2.3. Eller serbest YDT (A) anterior uzanma (B) posterolateral uzanma (C) posteromedial uzanma [54]. ....	15
Şekil 3.1. Denetimli öğrenmenin çalışma prensibi [56]. ....	21
Şekil 3.2. Denetimsiz öğrenmenin çalışma prensibi [56]. ....	22
Şekil 3.3. Yarı-denetimli öğrenmenin çalışma prensibi [56]. ....	22
Şekil 3.4. Takviyeli öğrenmenin çalışma prensibi [56]. ....	23
Şekil 3.5. Karar ağaçları regresyon algoritmasının örnek çıktısı [61]. ....	26
Şekil 3.6. Rastgele orman regresyon algoritmasının çalışma prensibi [63]. ....	27
Şekil 3.7. Gradyan artırma regresyon algoritmasının çalışma prensibi [65]. ....	29
Şekil 4.1. Gerekli kütüphanelerin import edilmesi. ....	32
Şekil 4.2. Veri setinin örnek içeriği. ....	33
Şekil 4.3. Kategorik verilerin tespit edilmesi. ....	34
Şekil 5.1. SAĞANTR% değeri tahmininde model performansları. ....	41
Şekil 5.2. SOLANT% değeri tahmininde model performansları. ....	42
Şekil 5.3. SAĞPM% değeri tahmininde model performansları. ....	43
Şekil 5.4. SOLPM% değeri tahmininde model performansları. ....	45
Şekil 5.5. SAĞPL% değeri tahmininde model performansları. ....	47
Şekil 5.6. SOLPL% d değeri tahmininde model performansları. ....	48
Şekil 5.7. TOPLAMORT% değeri tahmininde model performansları. ....	50

## ÇİZELGELER DİZİNİ

### Sayfa

Çizelge 2.1. Veri setinde bulunan bireylerin kişisel ve fiziksel özellikleri.....	17
Çizelge 2.2. Veri setinde bulunan YDT verilerinin örneği. ....	18
Çizelge 4.1. One-Hot kodlama uygulanmış veriler.....	36
Çizelge 5.1. SAĞANTR% değeri tahmininde regresyon algoritmaları sonuçları. ....	40
Çizelge 5.2. SOLANTR% değeri tahmininde regresyon algoritmaları sonuçları.....	42
Çizelge 5.3. SAĞPM% değeri tahmininde regresyon algoritmaları sonuçları. ....	43
Çizelge 5.4. SOLPM% değeri tahmininde regresyon algoritmaları sonuçları.....	44
Çizelge 5.5. SAĞPL% değeri tahmininde regresyon algoritmaları sonuçları. ....	46
Çizelge 5.6. SOLPL% değeri tahmininde regresyon algoritmaları sonuçları.....	47
Çizelge 5.7. Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş değişkenleri kullanarak TOPLAMORT% değerinin regresyon algoritmaları sonuçları. ....	49
Çizelge 5.8. Cinsiyet, Yaş, BMI, Antrenman Yılı ve Branş değerlerine ile YDT verisi (SAĞANTR%) verilerek geri kalan YDT değerlerinin tahmini. ....	51

## SİMGELER VE KISALTMALAR DİZİNİ

### SİMGELER

- $y$  : tahmin edilmeye çalışılan değer  
 $b_n$  : bağımsız değişkenlerin katsayıları  
 $x_n$  : bağımsız değişkenlerin değerleri  
 $X_{sc}$  : ölçeklendirilmiş veri kümesi  
 $X_{min}$  : veri kümesindeki en küçük değer  
 $X_{max}$  : veri kümesindeki en büyük değer  
 $\sigma$  : standard sapma  
 $z$  : z skoru  
 $\mu$  : ortalama değer  
 $\sum n$  : n değeri için toplam sembolü

## KISALTMALAR

YDT	: Y-Denge Testi
SEBT	: Star Excursion Balance Test (Yıldız Gezi Denge Testi)
IMU	: Inertial Measurement Unit (İnersiyal Ölçüm Ünitesi)
ANT	: Anterior
PM	: Postreromedial
PL	: Posterolateral
CSV	: Comma Separated Values
GRNN	: Generalization Regression Neural Network
MSE	: Mean Square Error (Ortalama Kare Hatası)
RMSE	: Root Mean Square Error (Kök Ortalama Kare Hatası)
R2	: R-Kare
MLR	: Multiple Linear Regression (Çok Değişkenli Doğrusal Regresyon)
BMI	: Body Mass Index (Vücut Kitle İndeksi)
RNN	: Recurrent Neural Network (Özyinemeli Sinir Ağı)
XGBoost	: Extreme Gradient Boosting (Aşırı Gradyan Artırma)
AdaBoost	: Adaptive Boosting (Adaptif Artırma)
SGD	: Stochastic Gradient Descent (Stokastik Gradyan İniş)

## BÖLÜM 1

### GİRİŞ

Spor bilimlerinde, atletlerin denge ve hareket kabiliyetlerini objektif bir şekilde değerlendirmek, performanslarını artırmak ve yaralanma risklerini azaltmak büyük bir öneme sahiptir. Bu bağlamda, Y-Denge Testi (YDT) yaygın bir yöntem olarak kullanılmaktadır. Ancak, YDT'nin uygulanmasında ve sonuçlarının yorumlanmasında bir dizi zorlukla karşılaşılmaktadır. Bu zorluklar, geçerli sonuçların elde edilmesini bilimsel açıdan engelleyebilir.

YDT'nin etkili ve güvenilir sonuçlar vermesi, doğru uygulama ve eğitim gerektirir. Testi uygulayan kişinin yeterli eğitim ve deneyime sahip olması, sporcuların doğru hareketleri yapmalarını sağlamak ve sonuçların güvenilirliğini artırmak için önemlidir. Bununla birlikte, her sporcu farklı denge yeteneklerine sahip olabilir ve yaş, cinsiyet, antrenman geçmişi gibi faktörler sonuçları etkileyebilir. Bu bireysel farklılıkların göz önünde bulundurulması, elde edilen sonuçların anlamlı ve doğru bir şekilde yorumlanabilmesi için gereklidir.

Ölçümlerin hassas ve tekrarlanabilir olması, YDT'nin güvenilir sonuçlar üretmesi için önemlidir. Ölçüm cihazlarının doğru kalibrasyonu ve standart test koşullarının sağlanması gereklidir. Ayrıca, elde edilen verilerin doğru bir şekilde analiz edilmesi ve yorumlanması da deneyimli bir uzmanlık gerektirir. Bu aşamada, hangi yönde ne kadar uzandığının ve denge yeteneğinin nasıl değerlendirileceğinin belirlenmesi önemlidir.

Zaman kısıtlamaları, YDT'nin hızlı ve etkili bir şekilde yapılabilmesi için yeterli zamanın olması gerekliliğini ortaya koymaktadır. Ancak, sınırlı zaman içinde detaylı ve doğru sonuçlar elde etmek zor olabilir. Bu nedenle, testin hızlı ve etkin bir şekilde gerçekleştirilebilmesi için uygun zaman yönetiminin sağlanması önemlidir. Aynı



zamanda, testin yapıldığı zemin yüzeyi gibi çevresel faktörlerden etkilenme riski de vardır. Bu nedenle, standart bir test ortamının sağlanması ve çevresel faktörlerin mümkün olduğunca kontrol altında tutulması gereklidir. Ayrıca, sporcuların motivasyon seviyeleri ve fiziksel yorgunluğu da YDT sonuçlarını etkileyebilir. Bu faktörlerin sporcuların en iyi durumda olmalarının sağlanması ile kontrol altında tutulması gereklidir. Ancak, YDT verilerinin analizi ve yorumlanması karmaşık olabilir ve deneyimli uzmanlık gerektirebilir. İşte bu noktada, makine öğrenimi regresyon algoritmalarının kullanımı, spor biliminin bu önemli alanında yeni bir perspektif sunabilir.

Makine öğrenimi regresyon algoritmaları, büyük ve karmaşık veri setlerini analiz etme yeteneği ile bilimsel araştırmalara büyük katkı sağlayabilir. YDT verilerinin geniş bir örnekleme toplandığı düşünüldüğünde, bu verilerin makine öğrenimi algoritmaları kullanılarak tahmin edilmesi, atletlerin denge performansını daha hassas bir şekilde değerlendirmeyi mümkün kılabilir. Bu tahminler, atletlerin spora geri dönüş süreçlerinde, yaralanma risklerinin değerlendirilmesinde ve antrenman programlarının tasarlanmasında rehberlik edebilir. Ayrıca, makine öğrenimi algoritmaları, YDT verilerinin karmaşıklığını anlama ve farklı faktörler arasındaki ilişkileri tanımlama yeteneği ile önemli bir avantaj sunar. Bu algoritmalar sayesinde, yaş, cinsiyet, antrenman geçmişi gibi etkenlerin denge performansı üzerindeki etkisi daha iyi anlaşılabilir. Böylece, sporda bireyselleştirilmiş yaklaşımların geliştirilmesi ve atletlerin özel ihtiyaçlarına uygun programların oluşturulması mümkün hale gelir. Literatürde yapılan çalışmalar incelendiğinde makine öğreniminin spor bilimlerinde önemli bir yere sahip olduğu görülmektedir.

İncetaş vd. [1] tarafından yapılan bir araştırmada, erkek atletlerin uzun atlama mesafesini tahmin etmek için en etkili makine öğrenme algoritmasının belirlenmesi amaçlanmıştır. Araştırma kapsamında yaş ve hız değişkenleri kullanılarak 328 geçerli atlama verisi toplanmıştır. Bu veriler 73 Türk erkek atlete aittir. Analiz sonuçları, uzun atlama performansının seçilen bağımsız değişkenler ile belirlenebileceğini göstermiştir. 5-kat çapraz doğrulama tekniği ile modellerin performansı değerlendirilmiş ve Gradyan Artırma Regresyon Ağaçları algoritmasının en iyi sonucu 0.0865 Ortalama Kare Hatası (MSE) değeri ile elde ettiği belirlenmiştir. Bu çalışma,

antrenörlerin erkek atletlerin uzun atlama performansını tahmin etmek için önerilen makine öğrenme yaklaşımını kullanabileceği sonucuna varmıştır.

Son araştırmalarda, YDT performansını tahmin etmek için geliştirilmiş regresyon sinir ağları ve derin öğrenme yaklaşımları gibi çeşitli teknikler kullanılmıştır. Akay vd. [2] tarafından yapılan bir çalışmada, çocukların durarak uzun atlama mesafesini tahmin etmeye yönelik yeni bir yaklaşım geliştirilmiştir. Bu çalışma, 6 ile 13 yaş aralığındaki çocukları kapsayarak Genelleştirilmiş Regresyon Sinir Ağı (GRNN) kullanmış ve cinsiyet, yaş, ağırlık, boy, antrenman yılı, bacak uzunluğu ve spor branşı gibi önemli değişkenlerin tahmindeki etkisini incelemiştir. Yapılan analizler, 10-kat çapraz doğrulama ve Kök Ortalama Kare Hatası (RMSE) değerlerini içermiştir. Ayrıca, geleneksel Çoklu Doğrusal Regresyon (MLR) modelleri ile karşılaştırma da yapılmıştır. Çalışmanın sonuçları, GRNN tabanlı tahmin modellerinin MLR ile geliştirilen modellere kıyasla daha yüksek bir tahmin doğruluğuna sahip olduğunu göstermektedir. Bu çalışma, çocukların spor performansını tahmin etmeye yönelik literatüre önemli bir katkı sağlamaktadır.

Gil-Martin vd. [3] tarafından yapılan çalışmada, YDT olarak bilinen dinamik denge değerlendirmesinin otomatik puanlaması için derin öğrenme yaklaşımının önerildiği görülmektedir. Bu yöntem, hareket sırasında inersiyel sinyalleri kaydetmek amacıyla giyilebilir bir sensör kullanarak normalize edilmiş ulaşma mesafesini tahmin ederek YDT'nin puanlamasını otomatik olarak gerçekleştirmeyi hedeflemektedir. Araştırmada, derin sinir ağına uygun veri beslemek için çeşitli sinyal işleme tekniklerini değerlendirmiş ve bu değerlendirme, özyinelemeli sinir ağları (RNN) temel alınarak gerçekleştirilmiştir. Önerilen derin sinir ağı yapısı, zamansal desenleri modellemek için uzun kısa dönem bellek katmanlarını içerir ve bacak uzunluğuna göre normalize edilmiş tahminini hesaplamak için ek bir tam bağlı katmanı bulundurur. Araştırmada 407 katılımcının YDT değerlendirmelerini içeren bir veri kümesi kullanılarak yapılan analizler, derin öğrenme yaklaşımının YDT puanlamasında başarılı sonuçlar elde etmekte olduğunu göstermektedir. Bu çalışma, farklı spor dallarından genç ve orta yaşlı katılımcılardan oluşan kapsamlı bir veri seti kullanılarak gerçekleştirilmiş ve derin öğrenme yaklaşımının dinamik denge değerlendirmesinde kullanılabilirliği üzerine önemli bir katkı sunmuştur.

Johnston vd. [4] tarafından gerçekleştirilen bir çalışmada, YDT'nin dinamik denge değerlendirmeleri için yaygın olarak kullanılan bir yöntem olduğu belirtilmektedir. Bu araştırma, bir lomber inersiyel ölçüm ünitesinin (IMU) tek başına üç farklı YDT hareket yönü arasındaki farkı ve yorgunluk öncesi ve sonrası denge performansını ayırt etme yeteneğini incelemeyi amaçlamaktadır. Çalışmada on beş katılımcıya lomber IMU takılmış ve YDT hareketleri yorgunluk öncesi ve sonrasında gerçekleştirilmiştir. Elde edilen IMU verileri kullanılarak rastgele orman sınıflandırıcıları eğitilmiş ve sonuçlar değerlendirilmiştir. Araştırma sonuçları, lomber IMU'nun farklı YDT hareket yönlerini doğru bir şekilde ayırt edebildiğini ve yorgunluk öncesi ve sonrası denge performansını orta düzeyde doğrulukla sınıflandırabildiğini göstermektedir. Bu bulgular, makine öğrenimi yöntemlerinin klinik denge değerlendirmeleri üzerindeki etkisini vurgulayarak, sporcuların denge performansını daha hassas bir şekilde değerlendirmenin ve yorgunluk gibi faktörlerin etkisini anlamının yeni bir yaklaşım sunabileceğine işaret etmiştir.

Spor alanında Uslu vd. [5] tarafından yapılan bir çalışmada, kuvvet-hız profili testinin sporcuların performansını ve uygun antrenman programlarını belirlemede yaygın olarak kullanıldığını ancak test protokolünün yüksek ağırlıklar ve tekrarlar nedeniyle sıkıntılara yol açtığını göstermektedir. Bu nedenle, farklı ağırlıklardaki sıçrama yüksekliklerinin tek tekrarlı ölçüm verisi kullanarak makine öğrenme modelleri ile tahmin edilmesi amaçlanmıştır. Akdeniz Üniversitesi Spor Bilimleri Fakültesi'ndeki 52 sporcuyla yapılan çalışmada, katılımcıların demografik özellikleri ölçülmüş ve farklı ağırlıklarda sıçrama yükseklikleri belirlenmiştir. Elde edilen veriler normalize edilmiş ve makine öğrenme modelleri kullanılarak dikey sıçrama yükseklikleri tahmin edilmiştir. Beş farklı model arasından Gaussian Süreç Regresyonu'nun en yüksek başarıyla dikey sıçrama yüksekliği tahmini yaptığı görülmüştür.

Perez vd. [6] tarafından yapılan bir çalışmada, sporcuların alt ekstremite temas dışı yumuşak doku yaralanma riskinin önceden tahmin edilmesi amacıyla makine öğrenimi teknikleri kullanılmıştır. Makine öğrenimi, bilgisayar sistemlerine veri analizi ve öğrenme yeteneği kazandıran bir alanı ifade eder. Bu çalışmada kullanılan makine öğrenimi yöntemleri, sezon öncesi alınan anketler ve saha tabanlı testlerden elde edilen

verileri işleyerek sporcuların yaralanma riskini tahmin etmek için etkili bir şekilde kullanılmıştır.

Oytun vd. [7] tarafından yapılan bir çalışmada kadın hentbol oyuncularının spor performansını tahmin etmek için farklı makine öğrenme modellerini değerlendirmekte ve üstün modeli kullanarak tahmin edilen performansı etkileyen faktörleri belirlemektedir. Lineer regresyon, karar ağacı, destek vektör regresyonu gibi modeller kullanılarak serbest eller ve eller kalçada karşı zıplama, sprint, mekik koşusu ve çeviklik performansı tahmin edilmiştir. Radyal taban fonksiyonlu sinir ağı modeli, diğerlerine göre daha üstün performans sergilemiş ve çeşitli spor performans türlerini yüksek doğrulukla tahmin etme yeteneği göstermiştir. Bu çalışma, makine öğrenme yöntemlerinin hentbol oyuncularının performans tahmini alanında potansiyelini vurgulamaktadır.

Jauhiainen vd. [8] tarafından yapılan bir çalışmada, tahmin edici makine öğrenme yöntemlerinin veri odaklı bir şekilde spor yaralanma risk faktörlerini tespit etmede nasıl kullanılabileceğini göstermektedir. Üç boyutlu hareket analizi ve fiziksel veriler kullanılarak yapılan çalışmada, 314 genç basketbol ve floorbol oyuncusu üzerinde orta düzeyde ve ciddi diz ve ayak yaralanmalarını tahmin etmek amacıyla doğrusal ve doğrusal olmayan yöntemler uygulanmıştır. Sonuçlar, rastgele orman ve lojistik regresyon modelleri ile belirlenen on iki tutarlı ve yirmi tahminci risk faktörünün yaralanma tahmininde kullanılabileceğini göstermektedir. Önerilen modellerin tahmin gücü düşük olmasına rağmen, belirli tahmin edici yaralanma risk faktörlerinin tespit edilebileceği sonucuna işaret etmektedir.

Yuanqi vd. [9] tarafından gerçekleştirilen bir çalışma, sporcuların fiziksel kondisyon testleri ile yaralanma riski arasındaki ilişkiyi önceden tahmin etmek için makine öğrenimi yöntemlerini kullanmıştır. Bu çalışma, 17 genç kadın basketbol oyuncusunun verilerini kullanarak maliyet duyarlı sinir ağı tabanlı bir tahmin modeli oluşturmuş ve bu modelin yaralanma riskini başarıyla öngördüğünü göstermiştir. Elde edilen sonuçlar, modelin iyi bir ayırım ve kalibrasyon sağladığını ortaya koymaktadır, bu da eğitim pratiğinde kullanım potansiyeline sahip olduğunu vurgulamaktadır.

Şimşek vd. [10] tarafından yapılan çalışmada ise yalnızca fizyolojik parametreler kullanılarak bocce sporcularının performans sınıfları yüksek doğrulukla tahmin edilmiş ve bu yeni bulgularla literatüre katkı sağlanmıştır. Destek Vektör Makineleri-Radyal Tabanlı Fonksiyon çekirdeği, yüksek performanslı bocce oyuncularını ve düşük performanslı bocce oyuncularının %75'ni doğru bir şekilde tahmin etmiştir. Makine öğrenimi kullanarak denge verilerinden spor performansını tahmin etmek, yüksek potansiyelli bocce sporcularını seçmek için zaman tasarrufu sağlayan bir yaklaşım olarak gösterilmiştir.

Roobles Palazon vd. [11]. yaptıkları çalışmada, elit olmayan genç erkek futbol oyuncularında alt ekstremitte yumuşak doku yaralanmalarını önceden tahmin etmek amacıyla makine öğrenimi modelleri oluşturmayı ve bu modellerin performansını karşılaştırmayı amaçlamaktadır. İspanyol amatör kulüplerinin akademilerinden 260 oyuncuyla gerçekleştirilen çalışma, kişisel özellikler, psikolojik yapılar ve fiziksel ölçümleri içeren bir değerlendirme ile başlamış, ardından bir rekabet sezonu boyunca alt ekstremitte yaralanmaları izlenmiştir. Çalışma, altı alan tabanlı ölçümün alt kümesini kullanarak en iyi uyumlu tarama modelinin yaralanma riskini başarılı bir şekilde tahmin ettiğini ve bu modelin genç futboldaki yaralanma yönetiminde kullanılabileceğini göstermektedir.

Manshadi vd. [12] tarafından yapılan bir çalışmada, denge testi puanlarına göre katılımcıları sınıflandırmak ve dövüş sanatı sporcularını tespit etmek amaçlanmıştır. Çalışmada, statik ve dinamik denge indeksleri, dört farklı testten elde edilmiştir. İran'daki bir ortaokulda gerçekleştirilen araştırmaya elli sağlıklı gönüllü katılmıştır. Elde edilen denge testi puanları, destek vektör makinesi, k-en yakın komşular algoritması ve yapay sinir ağı gibi yöntemlerle sınıflandırıcılar için girdi olarak kullanılmıştır. Bu yaklaşım sonucunda %90,5 tespit doğruluğu elde edilmiştir. Elde edilen sonuçlar, denge indekslerinin dövüş sanatı sporcularını tespit etmek için etkili bir gösterge olduğunu ve aynı zamanda yetenek tanıma amacıyla da kullanılabileceğini ortaya koymaktadır.

Makine öğrenimi regresyon algoritmalarının YDT verilerinin tahmininde spor bilimine sağlayabileceği potansiyel büyük öneme sahiptir. Bu algoritmalar sayesinde,

atletlerin denge performansı daha kesin ve objektif bir şekilde değerlendirilebilir, yaralanma riskleri daha etkili bir şekilde yönetilebilir ve antrenman programları daha verimli bir şekilde tasarlanabilir. Bu da sporcuların performansını artırmaya ve spor bilimine yeni bir boyut kazandırmaya olanak tanır. Çalışmanın amacı, YDT'nin karşılaştığı bu zorluklara dikkat çekmek ve makine öğrenimi yöntemlerinin kullanımının, bu zorlukların üstesinden gelmede nasıl yardımcı olabileceğini incelemektir. Bu sayede, sporcuların denge performanslarının daha hassas bir şekilde değerlendirilmesine ve antrenman programlarının daha etkili bir şekilde tasarlanmasına olanak tanıyan yeni bir yaklaşım sunmayı amaçlamaktadır.

Çalışma toplamda altı ana bölümden oluşmaktadır. İlk bölüm olan “Giriş” kısmıyla başlayarak, tezin genel amaçları ve araştırma soruları tanıtılmıştır. İkinci bölümde, “Denge ve Y-Denge Testi” incelenmiştir. Bu bölümde denge kavramı ve Y-Denge testinin temel prensipleri ayrıntılı olarak ele alınmıştır. Üçüncü bölümde, “Makine Öğrenimi ve Regresyon” başlığı altında makine öğreniminin temel ilkeleri ve regresyon algoritmaları açıklanmıştır. Bu bölümde, karar ağaçları, rastgele orman, adaptif artırma, gradyan artırma gibi regresyon yöntemleri detaylarıyla anlatılmıştır. Dördüncü bölümde, “Y-Denge Testi Verilerinin Makine Öğrenimi Algoritmaları ile Analizi” başlığı altında veri ön işleme adımları ele alınmıştır. Eksik verilerin işlenmesi, normalizasyon ve kategorik verilerin işlenmesi gibi adımlar detaylı olarak açıklanmıştır. Ayrıca, model seçimi süreci bu bölümde detaylarıyla incelenmiştir. Beşinci bölümde, “Deneysel Çalışmalar” başlığı altında gerçekleştirilen deneyler ve elde edilen sonuçlar paylaşılmıştır. Bu bölümde, tezin temel amacını desteklemek için yapılan pratik çalışmalar ve elde edilen veriler açıklanmıştır. Altıncı bölümde ise, “Sonuçlar ve Öneriler” kısmında elde edilen sonuçlar değerlendirilmiş ve tezin önerdiği yaklaşımların genel etkisi tartışılmıştır. Son olarak, “Kaynaklar” bölümünde tez çalışmasının derlenmesinde kullanılan kaynaklar ve referanslar sunulmuştur.

Bu altı ana bölüm, tez çalışmanızın yapısal temelini oluşturur ve her bir bölüm, araştırmanızın farklı yönlerini ele alarak konuyu kapsamlı bir şekilde incelemektedir.

## BÖLÜM 2

### DENGE VE Y-DENGE TESTİ

Günümüz spor alanlarında, sporcuların performanslarını maksimum düzeyde optimize etmek ve aynı zamanda sakatlanma riskini en aza indirmek temel hedefler arasında öncelikli bir yer tutmaktadır. Sporun rekabetçi doğası ve sporcuların üst düzey performans gösterme isteği, antrenörleri, fizyoterapistleri ve spor hekimlerini sporcuların sağlığını ve performansını en iyi şekilde yönetmeye yönlendirmiştir. Bu bağlamda, denge yeteneği, sporcuların başarısı ve uzun vadeli spor kariyerleri için belirleyici bir faktör olarak ön plana çıkmaktadır.

Denge yeteneğinin sporcular için taşıdığı önem hem spor performansı hem de yaralanma riski üzerinde etkili bir rol oynamaktadır. Sporcuların oyun içerisindeki çeşitli hareketleri gerçekleştirirken dengeli bir şekilde durma, hızlı yönlendirme yapma ve ani değişikliklere uyum sağlama yetenekleri, genellikle performanslarını belirleyen faktörler arasında yer almaktadır. Ancak, yüksek performans hedeflerine ulaşma sürecinde sporcuların karşılaştığı yaralanma riski de göz ardı edilemez bir gerçektir. İşte bu noktada, denge yeteneği hem spor performansını artırmada hem de sakatlanma riskini azaltmada kritik bir role sahiptir.

Sporcuların bireyselleştirilmiş antrenman programlarının oluşturulması ve performanslarının artırılması, modern spor yönetiminin temel taşlarından birini oluşturur. Ancak bu amaçlara ulaşmak için, sporcuların fiziksel yeteneklerinin ve zayıf noktalarının objektif bir şekilde belirlenmesi gerekmektedir. İşte burada denge yeteneğinin ölçümü ve değerlendirmesi devreye girer. Denge yeteneğinin objektif bir şekilde değerlendirilmesi, sporcuların hangi alanlarda daha fazla çalışmalar gerektiğini anlamalarına ve antrenman programlarını buna göre şekillendirmelerine yardımcı olur. Bu sayede sporcular, performanslarını en üst düzeye çıkarmak için daha

etkili bir şekilde çalışabilirler. Bu noktada, Y-Denge Testi'nin spor ve denge arasındaki ilişkinin anlaşılması ve değerlendirilmesi açısından hayati bir rol oynadığı görülmektedir.

## **2.1. DENGE**

Denge, insan vücudunun istikrarını koruma yeteneğini ifade eden temel bir fizyolojik kavramdır. Spor alanında ve rehabilitasyon süreçlerinde denge, performansın optimize edilmesi ve sakatlık risklerinin azaltılması açısından kritik bir öneme sahiptir [13]. Genel olarak denge, kişinin hareket halindeyken veya hareketsiz durumlar karşısında ağırlık merkezini doğru pozisyonlayabilme yeteneğidir [14]. Denge, statik denge ve dinamik denge olarak iki temel alt başlık altında incelenir.

Statik denge, hareketsiz durumdayken vücudun dengeyi koruma yeteneğini ifade eder. Bu denge türünde, kişi dururken veya sabit bir pozisyonda iken vücut ağırlığını dengede tutar. Statik denge yeteneği, duruşun stabilitesini sağlama, vücudun yerçekimi etkisine karşı direnme ve istikrarı sürdürme kapasitesini içerir. Örnek olarak, bir kişinin tek ayak üzerinde durması, plank pozisyonunda durması veya yoga duruşları gibi hareketsiz durumlarda vücudun dengeyi sağlama kabiliyeti statik dengeyi örnekler. Bu denge türü, günlük yaşam aktiviteleri, spor performansı ve rehabilitasyon süreçleri gibi birçok alanda önemlidir. Statik denge yeteneği, sporcuların performansını etkileyebilir ve yaralanma riskini azaltmada da önemli bir rol oynar. Sporcuların statik denge yetenekleri, özellikle dengeli bir duruş veya pozisyon gerektiren spor dallarında, performanslarını ve etkinliklerini etkileyebilir.

Dinamik denge, sporcuların hareket halindeyken vücutlarını dengeleyerek stabiliteyi koruma yeteneğini ifade eder. Bu denge türü, kişinin hareket ederken, koşarken, zıplarken veya hızlıca yön değiştirirken vücut dengesini koruma yeteneğini içerir. Dinamik denge, kasların, sinir sisteminin ve iç kulak gibi denge ile ilgili sistemlerin etkileşimiyle sağlanır. Sporcular için dinamik denge, özellikle hızlı hareketler ve ani yönlendirmeler gerektiren spor dallarında kritik bir öneme sahiptir. Örneğin, futbol, basketbol, tenis gibi sporlarda hızlı koşular, ani dönüşler ve yüksek hızda hareketler



sıkça gerçekleşir. Dinamik denge yeteneđi, sporcuların bu tür hareketleri düzgün ve etkili bir şekilde gerçekleştirebilmesi için gereklidir.

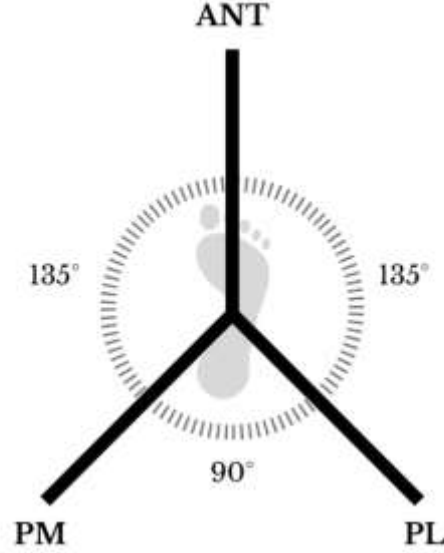
## 2.2. Y-DENGE TESTİ

YDT sporcuların dinamik denge ve fonksiyonel hareket yeteneklerini objektif bir şekilde değerlendiren önemli bir ölçüm yöntemidir. Bu test, vücut denge kontrolünü, hareket stabilitesini ve potansiyel sakatlık risklerini değerlendirme amacı güder [15]. Test, genellikle alt ekstremite fonksiyonel hareketlerini içerir ve sporcuların fiziksel yeteneklerini objektif bir şekilde değerlendirmek için yaygın olarak kabul görmüş bir yöntemdir [16]. YDT'nin sağladığı veriler, bireylerin denge performansını analiz etme ve spora özgü antrenman programlarını tasarlama konusunda koçlara ve spor eğitimlerine önemli bilgiler sunar.

Alt ekstremite nöromüsküler kontrolünün doğru bir şekilde değerlendirilmesini sağlamak için etkili ve klinik olarak uygulanabilir kabul edilen bir dinamik stabilite testidir [17]. Yıldız Gezi Denge Testini (SEBT) standart hale getirmek için geliştirilmiş, YDT; anterior (ANT), posterolateral (PL) ve posteromedial (PM) olmak üzere üç yönde uzanmayı içermektedir. SEBT'nin sekiz yönde nispeten daha zaman alıcı olmasından kaynaklı dinamik kapasitenin değerlendirilmesine yönelik testi yönetmede daha pratik ve daha kolay bir yol olarak ortaya çıkmıştır [18].

YDT kiti, ANT, PL ve PM pozisyonlarda düzenlenmiş üç adet PVC borunun bağlandığı bir platformdan oluşmaktadır. Boruların her birinin üzerine 0,5 cm aralıklarla bir ölçü yerleştirilmiştir. Şekil 2.1'de görüldüğü gibi arka yönlerin her biri ön yönden 135<sup>0</sup> ve aralarında 90<sup>0</sup> olacak şekilde konumlandırılmıştır. Test kitinin olmadığı durumlarda, aynı ölçümlerle yere sabitlenmiş mezuralarla uygulanabilir. Burada bir erişim göstergesi yoktur ve katılımcıdan uzanma ayağının parmak ucu ile en uzak noktaya dokunup çekmesi beklenir. Bu nedenle değerlendiricinin test esnasında daha fazla dikkat etmesi gerekmektedir. Elde edilen uzanma mesafe değerleri, özellikle sporcularda alt ekstremite yaralanmalarında yüksek risk taşıyan bireyleri tespit etmek ve rehabilitasyon sonrası dinamik fonksiyonel simetriyi

değerlendirmek için kullanılan önemli bir araç haline gelmiştir. Bu gelişme, spor alanında özellikle büyük bir öneme sahiptir. [19].



Şekil 2.1. Y-denge testi kitinin kurulumu.

Bir sporcu için işlevsel yeteneklerin önemli bir unsuru olan denge kabiliyeti, dik duruş sırasında vestibüler, proprioseptif ve görsel sistemlerden gelen afferent bilgilerin santral entegrasyonuna dayanır [13,20]. Dinamik denge, bireyin hareket sırasında kütleli merkezin toplam vücut stabilitesini koruma yeteneği olarak tanımlanır [21]. Temel motor becerilerinden biri olarak kabul edilen denge, fiziksel gelişim açısından önemli bir role sahiptir ve diğer motor yeteneklerin gelişimine katkı sağlar [22,23]. Spor performansı perspektifinden değerlendirildiğinde, hem statik hem de dinamik spor dallarında önemli etkileri gözlemlenir [24]. Ancak, kas-iskelet sistemi yaralanmaları, sporun doğal bir riski olarak karşımıza çıkar ve genellikle yarışma veya antrenman esnasında meydana gelir. Her yıl yaklaşık üç ile beş milyon spor kaynaklı yaralanmanın çoğunlukla alt ekstremitede gerçekleştiği tahmin edilmektedir [25,26]. Bu tür yaralanmaların, genellikle uzun vadeli sonuçlar olarak artan risk, zaman kaybı ve tıbbi maliyetler gibi olumsuz etkilere neden olduğu ifade edilmektedir [27]. Denge yetenek ile spor yaralanma riski arasındaki ilişki birçok vakada belirlenmiştir. Zayıf denge ile alt ekstremitte yaralanma riskinin arttığı gösterilmiştir [28]. Temas olmayan yaralanmalar oyun sırasında meydana gelen yaralanmaların yaklaşık %20'sini ve antrenman sırasında meydana gelen yaralanmaların yaklaşık %40'ını oluşturmaktadır

[26]. Bu nedenlerle, alt ekstremite denge ölçen tarama testlerinin kullanımı, spor hekimlerinin yaralanma riski taşıyan sporcuları tespit etmelerine yardımcı olabilir [29]. Ayrıca, ayak bileği ve diz yaralanmalarını önleme süreçlerinde denge testlerinin kritik bir rol oynadığı vurgulanmıştır. YDT'nin güvenliği çeşitli popülasyonlarda incelenmiş ve dinamik nöromusküler kontrolün güvenilir bir testi olarak orta ile yüksek kalitede kanıtlar sunulmuştur [30]. Testin, pre-adolesan, geç adolesan ve yetişkin sporcular için güvenilir bir araç olduğu bildirilmiştir. Ayrıca, erken adolesan dönemdeki kadınların dinamik dengeyi değerlendirmek için güvenilir bir araç olduğu bildirilmiştir [31]. Düzenli pilates egzersizi yapan kadınlar ile egzersiz yapmayan kadınların YDT performansları karşılaştırılmış; sonuçlar, düzenli pilates egzersizlerinin dinamik postüral kontrolü olumlu şekilde etkilediğini ve spor yapan kadınlarda uygulanan YDT'nin egzersiz programlarına katılmadan önce tarama ve yaralanma tahmini amacıyla da kullanılabileceğini öne sürmektedir [32]. Diğer çalışmalarda limb (uzuv), cinsiyet ve antropometrik özellikler gibi faktörlerin YDT performansı üzerindeki etkisi incelenmiştir. Cinsiyet ve ekstremite performansı karşılaştırmaları yapılması ve test performansı ile boy, vücut ağırlığı ve vücut kitle indeksi arasındaki ilişkilerinin incelenmesi amaçlanmıştır. Cinsiyetler arasında ortaya çıkan antropometrik ve performans ilişkileri arasındaki farklılıklardan dolayı tek bacaklı test performansını etkileyen faktörleri anlamak için daha fazla araştırmaya ihtiyaç duyulmaktadır [33]. Benzer şekilde, boy ve kilo gibi antropometrik özelliklerin geniş bir spor yelpazesi içinde yer alan çocukların YDT performansı ile ilişkili olduğu bulunmuştur [34].

### **2.2.1 Y-Denge Testi Uygulama Alanları**

Statik denge değerlendirmeleri önemli klinik bilgiler sunsa da, yaralanma durumlarında tek bacak üzerinde yapılan dinamik denge testleri, günlük yaşam aktivitelerini taklit ettikleri için tercih edilmektedir. Yaralanma sonrası kalıcı sorunlar ve artan yaralanma riski dinamik nöromusküler kontrol eksiklikleriyle ilişkilendirildiğinde, sporcularda bu boşlukları tespit etmek için YDT gibi tek bacaklı dinamik değerlendirmelerin kullanımını artmıştır [30]. Bu nedenlerle, yetişkinlerin spor ve rehabilitasyon süreçlerinde bu test, yaygın bir şekilde kullanılmaktadır [19,32].

Sporcuların fiziksel hazırlık ve performans değerlendirmelerinde önemli bir araç olarak kullanılan YDT, postüral kontrol ve dinamik denge performansını değerlendirmek açısından dikkate değerdir [35]. YDT'nin antrenmanlarla gelişen performansı, sporcuların spor katılımına hazır olma düzeyini değerlendirmekte kullanılabilmesi gibi, yaralanma riskini azaltma ve sporcuların performansını artırma potansiyelini vurgulamıştır [36]. Ayrıca, YDT'nin dinamik postüral kontrolü değerlendirilerek alt ekstremitelerde yaralanmalarını öngörme yeteneği birçok çalışmada vurgulanmıştır [37-39]. YDT'nin antrenmana dönüş süreçlerinde fonksiyonel sonuç ölçütü olarak kullanılabilirliği, özellikle ön çapraz bağ yaralanmalarında alt ekstremitelerde asimetrisini tespit etme ve antrenmanlara geri dönüşü etkileme potansiyelini yansıtmıştır [40,41]. Müsabaka seviyesine özgü dinamik denge normatif standartlarının, antrenmana dönüş sırasında fiziksel hazır olma durumunun değerlendirilmesinde önemli bir rol oynadığı belirtilmiştir [42]. Aynı zamanda, YDT'nin yaralanma rehabilitasyonunu takip etmek ve iyileşme sürecini değerlendirmek için de kullanılan güvenilir bir araç olduğu ifade edilmiştir [43]. Örneğin, ameliyat sonrası 6 ay içinde tek bacak dinamik denge bozukluklarının ölçülendirilmesine yardımcı olabilir [44]. Sporcuların fiziksel hazır olma durumunu ve potansiyel yaralanma riskini belirlemede önemli bir gösterge olarak YDT'nin kullanışlı bir araç olduğu görülmektedir [45,46]. Bu nedenle, YDT'nin sporcuların performansını değerlendirmek, yaralanma riskini belirlemek ve rehabilitasyon süreçlerini yönlendirmek için geniş bir uygulama potansiyeli taşıdığı söylenebilir [47].

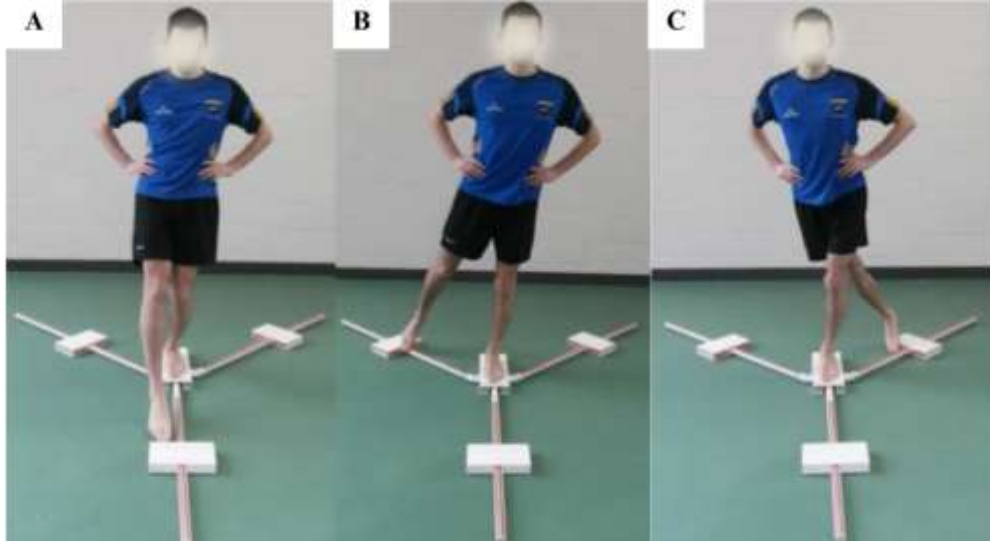
### **2.1.2. Y-Denge Testinin Uygulanması ve Ölçülmesi**

Y-Denge Testi'nin uygulaması, katılımcıların denge yeteneklerini değerlendirmek amacıyla yapılmaktadır. Test sırasında katılımcılar, YDT cihazında çıplak ayakla merkez plakada tek ayak üzerinde dururken, parmak ucu kullanarak uzanma göstergesini farklı pozisyonlarda (ANT, PL ve PM) iterek uzanma gerçekleştirirler. Uzanma mesafeleri, katılımcının merkezdeki ayak parmak ucundan ANT, ayak topuğundan ise PL ve PM pozisyonlarında uzanma göstergesine en yakın noktaya kadar olan mesafeler olarak santimetre cinsinden kaydedilir. Test öncesi, katılımcılara testin nasıl uygulanacağı konusunda kısa bir gösterim yapılır veya uygulama videosu izletilir [36].

Elde edilen verileri analiz etmek için, bacak uzunluğunun etkisini dengelemek amacıyla her yöndeki uzanma mesafeleri “(En İyi Uzanma Mesafesi / Bacak Uzunluğu) x 100” formülü ile normalize edilir. Normalize edilmiş ANT, PL ve PM puanları ortalaması alınarak toplam puan değeri hesaplanır. Ayrıca, bazı çalışmalarda bacak uzunluğu normalizasyonu olmadan da test sonuçları kullanılabilir [48].

Testin deneme ve ölçüm sayıları konusunda önceki çalışmalar, sonuçların daha istikrarlı olması için altı deneme ve üç ölçüm yapılmasını önermektedir [36]. Ancak farklı YDT versiyonları arasında anterior uzanma mesafesinde farklılıklar olduğu gözlemlenmiştir [49]. Bu nedenle, YDT'nin güvenilirliğini artırmak amacıyla özgül protokollerin geliştirilmesi önerilmiştir [12,50].

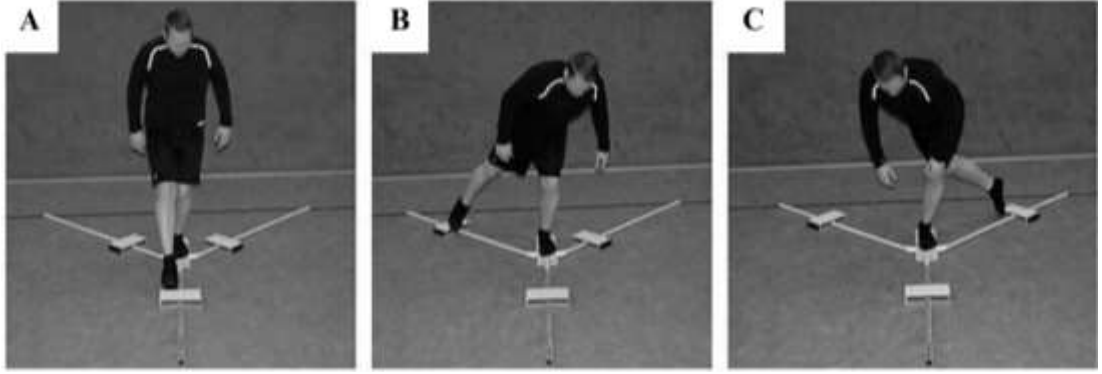
Genç sporcularda dokuz denemenin yapılmasını ve ölçülen üç denemenin ortalamasının analiz edilmesini önerirken, bazı çalışmalarda en iyi uzanma değerinin değerlendirmeye alındığı görülmektedir [51]. Genç sporcularda ek iki uygulama denemesinin gerekliliği, olgunlaşma aşamalarının farklılığına bağlanmıştır.



Şekil 2.2. Eller kalçada YDT (A) anterior uzanma (B) posterolateral uzanma (C) posteromedial uzanma [52].

YDT uygulamasında el pozisyonu için farklı protokoller kullanılabilir. Bu protokollerde, ellerin Şekil 2.2’de gösterildiği gibi kalçada olması veya Şekil 2.3’te gösterildiği gibi serbest hareket etmesi gibi farklı yaklaşımlar bulunmaktadır. Bu

farklılıkların sonuçları etkileyebileceği ve el pozisyonunun YDT puanlarını önemli ölçüde etkileyebileceği gözlemlenmiştir [53]. Ayrıca, ölçüm sırasında katılımcıların vücut ağırlığını uzanma ayağına aktarmaması, duruş ayağının topuğunu zeminden ayırmaması veya ellerini kalçadan ayırmaması gibi hataların tekrarlanması gerektiği belirtilmiştir.



Şekil 2.3. Eller serbest YDT (A) anterior uzanma (B) posterolateral uzanma (C) posteromedial uzanma [54].

### 2.3. Y-DENGE TESTİ VERİ KÜMESİ

Bu çalışmada, sporcuların dinamik denge yeteneklerinin analizi için Y-Denge Testi veri kümesi kullanılmıştır. Çalışmaya dahil edilme kriterleri, katılımcıların yaralanma öyküsüne sahip olmaması, katılımı engelleyecek bir sağlık sorununun bulunmaması, büyüme düzensizliği öyküsü olmamasını içermektedir. Bu kriterler, veri setinin güvenilirliğini ve analiz sonuçlarının geçerliliğini sağlamak amacıyla titizlikle belirlenmiştir.

Veri toplama süreci, katılımcılara önce test prosedürlerine alışmaları için bir alışma oturumu uygulanması ile başlamıştır. Ardından üç gün sonra, katılımcılar antropometrik özellikleri ve dinamik denge yeteneklerini değerlendirmek için bir test oturumuna katılmışlardır. Bu aşamadan sonra, katılımcılar ısınma için zıplama ve ayakta yürüyüş gibi hareketlerden oluşan bir ısınma oturumu gerçekleştirmişlerdir. Daha sonra, katılımcılar üç uzanma yönünde her bir bacakta altı deneme alıştırmaları yapmışlardır. Değerlendiriciler, katılımcılara istenen hareketi her iki bacak üzerinde üç uzanma yönünde uygulamalarını vurgulamış ve tek bacak üzerinde duruşun

önemini anlatmışlardır. Bu adımlar, veri setinin tutarlı ve güvenilir sonuçlar üretmesini sağlamak amacıyla gerçekleştirilmiştir.

Toplanan YDT verileri, sporcuların dinamik denge yeteneklerini yansıtan değerli bilgileri içermektedir. Bu veriler, çalışma kapsamında kullanılan makine öğrenme yöntemleri ile analiz edilerek farklı denge ölçütlerini anlamak ve sporcu performansını daha iyi değerlendirmek amacıyla kullanılmaktadır.

Veri etiketleme süreci, her katılımcının performansını belirli ölçütlere göre etiketlemeyi içermektedir. Bu etiketler, makine öğrenme algoritmalarının verileri analiz etmesi ve dengenin değişkenlerini anlaması için temel oluşturmaktadır. Bu şekilde, spor bilimlerinde YDT'nin veri odaklı analizi, sporcuların denge yeteneklerini daha iyi anlamamızı sağlayarak performansı artırmak ve sakatlanma riskini azaltmak gibi hedeflere katkıda bulunmaktadır.

Çalışma kapsamında kullanılan veri kümesi, Uşak ili sınırları içerisinde farklı spor branşlarında faaliyet gösteren spor kulüplerinden elde edilmiştir ve bu verilerin kullanımı için gerekli yasal izinler alınmıştır. Bu çalışma, 6-17 yaş arasındaki çocuk sporcuların dinamik denge verilerini içeren bir veri setini temel almaktadır. Toplamda 216 çocuk, farklı spor kulüplerinden ve yaş aralıklarından, bu araştırmaya dahil edilmiştir. Katılımcılar düzenli olarak futbol, basketbol, tenis, voleybol, hentbol, jimnastik, yüzme ve tekvando gibi sporlara katılımlarına göre 8 gruba ayrılmıştır. Her katılımcının kişisel bilgileri, branş bilgileri, boy ve kilo ölçümleri, antrenman yılı, baskın bacak bilgileri veri setine kaydedilmiştir. Üç dakikalık bir ısınma döneminden sonra, katılımcılar her bir bacaklarıyla üç farklı yatay yönde toplam altı deneme egzersizi gerçekleştirmiştir.

Pratik denemeler sırasında, araştırmacılar istenen hareketleri gerçekleştirirken (ön, posteromedial ve posterolateral yönlere doğru hareketler) tek bacak üzerinde dengelemenin önemi vurgulamıştır böylece her iki bacağın da ilgili yönlere ulaşabilmesi sağlanmaktadır. Bireysel bacak uzunluklarındaki varyasyonları hesaba katmak için anterior, posteromedial ve posterolateral değerler bacak uzunluğuna bölünerek normalize edilmiştir. Elde edilen bu normalize edilmiş değerler, her iki sağ

ve sol bacak için yüzde (%) olarak ifade edilerek SAĞANT%, SOLANT%, SAĞPM%, SOLPM%, SAĞPL% ve SOLPL% YDT değerleri veri setine eklenmiştir. Her iki bacağın normalize edilmiş ön değerlerinin Eşitlik 2.1’de verildiği gibi ortalaması ANT%ORT olarak adlandırılmıştır; aynı şekilde her iki bacağın normalize edilmiş posteromedial ve posterolateral değerlerinin ortalamaları Eşitlik 2.2 ve Eşitlik 2.3’deki gibi hesaplanarak PM%ORT ve PL%ORT olarak belirtilmiştir. Son olarak, ANT%ORT, PM%ORT ve PL%ORT Eşitlik 2.4’te gösterildiği gibi ortalamalarının hesaplanmasıyla TOPLAM%ORT elde edilerek veri setine kaydedilmiştir.

$$ANT\%ORT = \frac{SAĞANTR\% + SOLANTR\%}{2} \quad (2.1)$$

$$PM\%ORT = \frac{SAĞPM\% + SOLPM\%}{2} \quad (2.2)$$

$$PL\%ORT = \frac{SAĞPL\% + SOLPL\%}{2} \quad (2.3)$$

$$TOPLAM\%ORT = \frac{ANT\%ORT + PM\%ORT + PL\%ORT}{3} \quad (2.4)$$

Çizelge 2.1 ve Çizelge 2.2’de veri setinin bir kısmından örnekler gösterilmiştir.

Çizelge 2.1. Veri setinde bulunan bireylerin kişisel ve fiziksel özellikleri.

Cinsiyet	Branş	Yaş	Boy	Kilo	BMI	Antrenman Yaşı	Baskın Bacak
K	TENİS	7	125	25	16	2	1
E	BASKETBOL	9	138	30	15,753	3	1
K	CİM	11	138	29	15,228	6	1
E	FUTBOL	14	158	42,5	17,025	4	1
K	HENT	11	143	34	16,627	3	1
E	TEK	8	142	38	18,845	2	0
K	TENİS	6	120	20	13,889	1	1
K	VOL	9	154	42	17,710	2	1
K	YÜZME	14	161	55	21,218	5	0



Veri setinden verilen bu örnekte, her kişi için kaydedilen bilgiler mevcuttur. Cinsiyetler Kadın (K) ve Erkek (E) olarak sembolize edilmiştir. 8 farklı branştan örnekler gösterilmektedir. Baskın bacak bilgisi Sağ (1), Sol (0) olarak kodlanmıştır. Boy ve kilo bilgisi BMI'da, bacak uzunluğu YDT verilerinin normalize aşamasında, baskın bacak bilgisi ise kullanılmadığı için veri setinden temizlenmiştir. Burada kaydedilmiş YDT değerleri, bacak uzunluğuna bölünerek normalize edilmiş verilerdir.

Çizelge 2.2. Veri setinde bulunan YDT verilerinin örneği.

SAĞ ANTR %	SOL ANTR %	ANT% ORT	SAĞ PM%	SOL PM%	PM %ORT	SAĞ PL%	SOL PL%	PL% ORT	TOPLAM ORT%
60,847	55,556	58,201	127,513	119,048	123,280	107,937	81,481	94,709	92,063
49,367	44,304	46,835	82,700	100,422	91,561	78,481	70,886	74,684	71,027
65,432	65,021	65,226	92,181	93,416	92,798	84,362	86,008	85,185	81,070
65,201	64,835	65,018	102,930	100,366	101,648	103,297	106,593	104,945	90,537
61,404	57,018	59,211	92,982	92,544	92,763	83,772	82,456	83,114	78,363
62,222	65,778	64,000	102,222	104,889	103,556	107,556	109,333	108,444	92,000
52,688	50,538	51,613	103,226	93,011	98,118	61,828	63,978	62,903	70,878
53,778	53,778	53,778	94,667	90,667	92,667	71,111	56,444	63,778	70,074
59,259	64,198	61,728	104,115	106,996	105,556	94,650	98,765	96,708	87,997

Makine öğrenimi algoritmaları kullanarak regresyon problemlerinde verilerin etiketlenmesi, modelin doğru ve güvenilir tahminler yapabilmesi için hayati bir öneme sahiptir. Etiketlenmiş veriler, modelin girdi verileriyle çıktı değerleri (etiketler) arasındaki ilişkiyi öğrenmesine yardımcı olmaktadır.

Toplanan YDT verileri ve kişisel bilgiler bir Excel dosyasında tutulmuştur. Veri kümesini veri paylaşımı, işbirliği, analiz ve hafiflik açısından birçok avantaj sunduğu için CSV (Comma-Separated Values) formatına dönüştürülmüştür. Tez çalışmasında kullanılacak olan makine öğrenimi algoritmalarının YDT alanında kullanıldığı çalışmalar da incelenerek kapsamlı bir literatür taraması yapılmıştır.

YDT sonuçlarını tahmin etmek için makine öğrenme yöntemlerinin kullanılma fikri, testin duyarlılığı ve doğruluğu açısından önemlidir, çünkü hataları önlemeye ve test sonuçlarını etkileyen faktörleri analiz ederek sporcuları uygun tedaviye yönlendirmede yardımcı olabilir. Son yıllarda, makine öğrenme teknikleri tıbbi alanda yaygın bir şekilde kullanılmaktadır. Makine öğrenme, verileri matematiksel modeller

kullanarak öğrenerek belirli görevlerde performansı artırmaktadır. Bu algoritmalar, verilerdeki desenleri ve ilişkileri tanır ve bu bilgileri yeni veriler için tahminlerde bulunmak veya kararlar almada kullanır. YDT sonuçları etkileyen değişken faktörlerin bulunduğu bir denge testidir. Bu faktörler yaş, cinsiyet, boy, kilo ve antrenman yaşını içerir.

Bu tez çalışması, YDT sırasında spor alanında yer alan çocukların performansını belirleyen faktörleri analiz etmek için makine öğrenme algoritmalarının kullanımını tartışmaktadır. Çalışmamızda, yaş, cinsiyet, antrenman yaşını, vücut kitle indeksini (BMI) ve sporcuların branş bilgisini içeren değişkenlerin YDT sonuçları üzerindeki etkileri incelenmektedir. Amacımız, YDT sonuçlarından elde edilen verileri kullanarak çocukların dinamik denge yeteneklerini etkileyen faktörleri belirleyebilecek bir yöntem geliştirmek ve bu faktörlere dayalı kişiselleştirilmiş eğitim programları oluşturmaktır. Ayrıca, bu çalışmada farklı regresyon algoritmalarının performansı değerlendirilmekte ve sonuçlar MSE, R2 ve RMSE gibi kriterlerle karşılaştırılmaktadır. Sonuç olarak, hangi regresyon algoritmasının spor alanında yer alan çocukların denge yeteneklerini daha iyi tahmin ettiği belirlenmektedir. Bu bulgular, çocukların daha etkili bir şekilde eğitim programları oluşturmalarına ve dinamik denge yeteneklerini geliştirmelerine olanak tanımaktadır.

## BÖLÜM 3

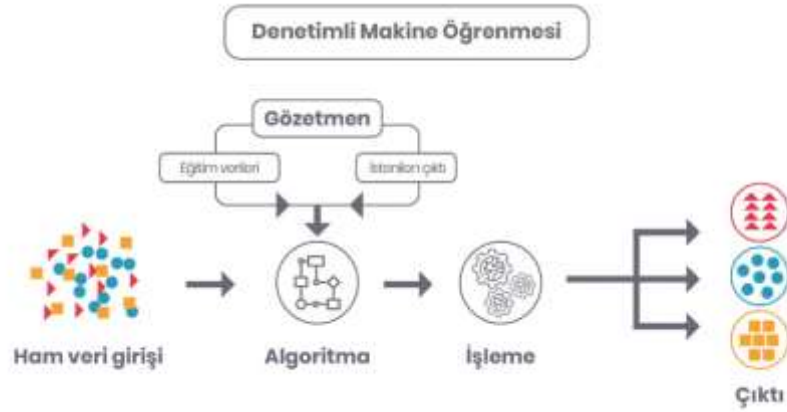
### MAKİNE ÖĞRENİMİ VE REGRESYON

Makine öğrenimi alanında regresyon analizleri, karmaşık ilişkileri anlamak ve gelecekteki değerleri tahmin etmek için yaygın olarak kullanılan güçlü tekniklerdir. Bu çalışma, sporcuların denge performansını y-denge testi sonuçları üzerinden tahmin etmeyi amaçlamaktadır. Her bir denge yönde (sağ anterior, sol anterior, sağ posteromedial, sol posteromedial, sağ posteriolateral ve sol posteriolateral) ayrı ayrı regresyon modelleri geliştirilmiştir. Bu modeller, sporcuların yaş, cinsiyet, vücut kitle indeksi ve antrenman yaşını içeren girdi verilerini kullanarak denge performansını tahmin etmek için tasarlanmıştır.

Her bir regresyon modeli, girdi verileri ile denge performansı arasındaki ilişkiyi matematiksel olarak ifade eder. Yaş, cinsiyet, vücut kitle indeksi ve antrenman yaşının denge performansı üzerindeki etkileri bu modeller aracılığıyla analiz edilmiştir. Bu analizler, regresyon denklemleri vasıtasıyla elde edilen tahminlerle gerçek denge performansı değerlerini karşılaştırarak yapılmıştır. Makine öğrenimi algoritmaları kullanılarak geliştirilen bu regresyon modelleri, sporcuların denge yeteneklerini değerlendirmek ve geliştirmek için değerli bir araç sunmaktadır. Ayrıca, bu modeller sporcuların yaralanma riskini tahmin etmekte de kullanılabilir ve sporculara özelleştirilmiş eğitim programları oluşturma konusunda rehberlik edebilir. Bu çalışmanın sonuçları, makine öğrenimi algoritmaları ve regresyon analizlerinin spor performansını anlamada ve tahmin etmede nasıl etkili bir rol oynayabileceğini göstermektedir.

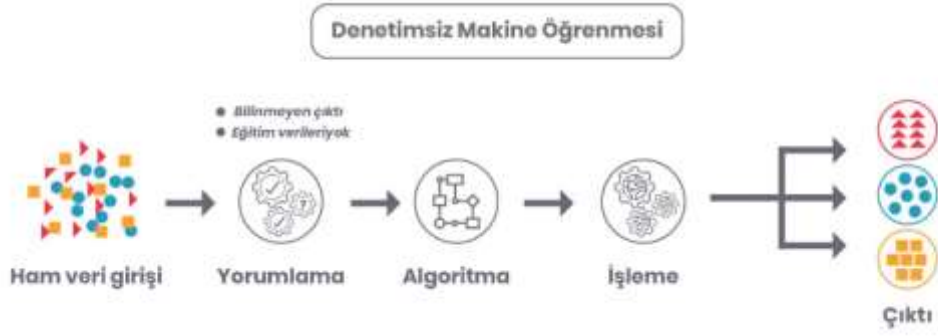
### 3.1. MAKİNE ÖĞRENİMİ

Makine öğrenimi, geniş veri kümelerinden bilgi çıkarmayı amaçlayan bir disiplindir. Denetimli öğrenme, denetimsiz öğrenme, yarı-denetimli, takviyeli öğrenme makine öğrenimi alanının temel yaklaşımlarını oluşturur. Denetimli öğrenme, algoritmaların belirli girdi verilerine dayalı olarak çıktı tahminleri yapmayı öğrendiği bir yaklaşımdır. Şekil 3.1’de denetimli öğrenmenin çalışma prensibine bir örnek verilmiştir. Eğitim verileri içerisinde girdi-veri çiftleri ve bu çiftlere ait hedef çıktı değerleri bulunur. Bu model, eğitildikten sonra yeni girdi verilerine ilişkin tahminlerde bulunabilir. Örneğin, sınıflandırma problemlerinde, bir algoritma doğru ve yanlış cevapları içeren bir veri kümesini kullanarak yeni verilerin doğru veya yanlış olma olasılığını tahmin edebilir. Ayrıca, regresyon problemlerinde ise sayısal çıktıları tahmin edebilir [55].



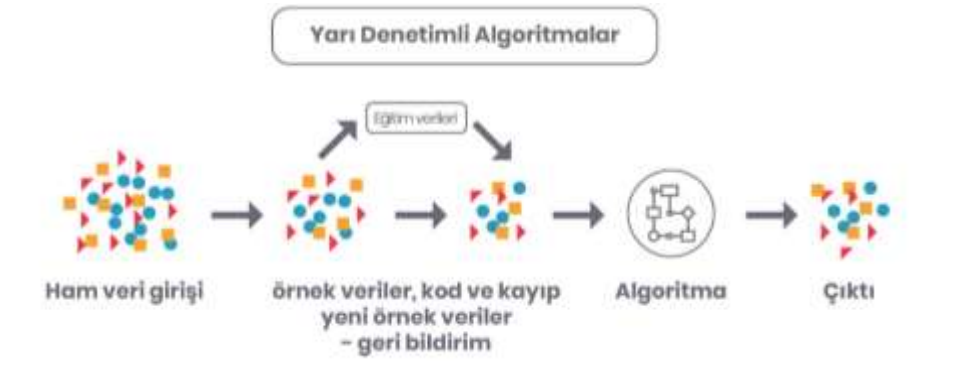
Şekil 3.1. Denetimli öğrenmenin çalışma prensibi [56].

Denetimsiz öğrenmede eğitim verileriyle ilişkili çıktılar sağlanmaz. Bu yaklaşımda, algoritmalar girdi verilerini analiz ederek içsel yapıları ve örüntüleri tanımaya çalışır. Şekil 3.2’de denetimsiz öğrenmenin çalışma prensibine bir örnek verilmiştir. Örneğin, kümelenme teknikleri ile benzer örnekler gruplandırılabilir veya boyut indirgeme yöntemleri ile veriler daha düşük boyutlarda temsil edilerek temel yapıları anlama çabası gösterilir. Denetimsiz öğrenme, veri keşfi, segmentasyon, örüntü tanıma gibi birçok alanda etkili bir şekilde kullanılır [57].



Şekil 3.2. Denetimsiz öğrenmenin çalışma prensibi [56].

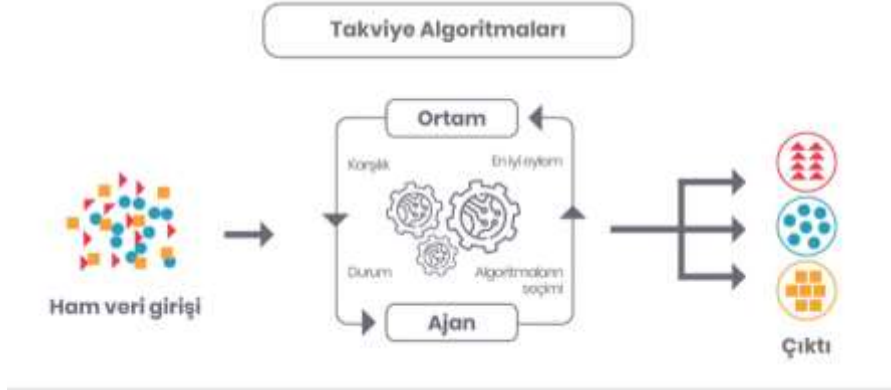
Yarı-denetimli öğrenme, etiketlenmiş ve etiketlenmemiş verilerin bir arada kullanıldığı bir öğrenme yaklaşımıdır. Şekil 3.3'te yarı-denetimli öğrenmenin çalışma prensibine bir örnek verilmiştir. Genellikle etiketlenmiş verilerin elde edilmesi zaman alıcı ve maliyetli olabilirken, etiketlenmemiş veriler daha kolay toplanabilir. Yarı-denetimli öğrenme, bu iki veri türünü birleştirerek daha geniş bir veri yelpazesi elde etmeyi amaçlar. Etiketli verilerin daha az olduğu durumlarda, etiketlenmiş verileri daha etkin bir şekilde kullanarak modelin performansını artırabilir. Özellikle sınıflandırma ve tahmin problemlerinde, yarı-denetimli öğrenme teknikleri, daha az etiketlenmiş veriyle daha iyi sonuçlar elde edilmesini sağlayabilir [58].



Şekil 3.3. Yarı-denetimli öğrenmenin çalışma prensibi [56].

Takviyeli öğrenme, bir ajanın bir çevre içinde deneyimler kazanarak en iyi eylemleri seçmeyi öğrenmeye çalıştığı bir öğrenme türüdür [59]. Şekil 3.4'te takviyeli öğrenmenin çalışma prensibine bir örnek verilmiştir. Ajan, çevre ile etkileşim halinde olup, her bir eylem sonucunda ödüller veya cezalar alır. Amacı, zaman içinde toplam

ödülü maksimize edecek eylemleri öğrenmektir. Bu tür öğrenme, özellikle karmaşık ve dinamik çevrelerde, insan müdahalesi olmaksızın öğrenme yeteneği kazandırabilir. Oyun oynama, robotik ve optimizasyon gibi alanlarda takviyeli öğrenme teknikleri önemli bir rol oynar [57].



Şekil 3.4. Takviyeli öğrenmenin çalışma prensibi [56].

Regresyon, denetimli öğrenim kategorisinde yer alırken, bir modelin girdi verilerine dayanarak sayısal çıktılar üretmesini sağlar. Bu sayede, veri setlerindeki ilişkileri anlamak ve gelecekteki değerleri tahmin etmek mümkün hale gelir. Makine öğreniminin ve özellikle regresyonun kullanımı, çeşitli disiplinlerdeki veri analiz ihtiyaçlarını karşılamada etkili bir araç olarak ortaya çıkmıştır.

### 3.2. MAKİNE ÖĞRENİMİ REGRESYON ALGORİTMALARI

Regresyon, istatistiksel analizler ve makine öğrenimi alanında sıklıkla kullanılan bir yöntemdir ve değişkenler arasındaki ilişkileri anlamayı, gelecekteki değerleri tahmin etmeyi ve veriler arasında bir tür fonksiyonel bağlantı kurmayı amaçlar. Temelde bir bağımlı değişkenin bir veya daha fazla bağımsız değişkene bağlı olarak nasıl değiştiğini incelemek için kullanılır. Regresyon analizi, bu değişkenler arasındaki matematiksel ilişkiyi tanımlamak için bir regresyon denklemi oluşturur. Bu denklem, bağımlı değişkenin değerini, bağımsız değişkenlerin değerlerine göre tahmin etmek için kullanılabilir. Regresyon analizleri geniş bir uygulama yelpazesine sahiptir ve özellikle tahmin yapma, ilişkileri anlama, risk değerlendirmesi, veri modelleme ve hipotez test etme gibi alanlarda yaygın olarak kullanılır. Bu analiz yöntemi,

istatistiksel çıkarımlar yapma yeteneği ve tahmin gücü nedeniyle hem akademik araştırmalarda hem de endüstriyel uygulamalarda değerlidir.

Regresyon, çeşitli algoritmalar, teknikler ve yaklaşımlarla gerçekleştirilebilir. Makine öğrenimi alanında özellikle çoklu regresyon, lojistik regresyon, destek vektör regresyonu, karar ağaçları ve derin öğrenme gibi farklı regresyon yöntemleri yaygın olarak kullanılmaktadır [53]. Makine öğrenimi içerisinde regresyon, özellikle sürekli değerlerin tahmin edilmesi gereken durumlarda kritik bir rol oynar. Regresyon algoritmaları, bir bağımlı değişkenin bir veya daha fazla bağımsız değişkenle ilişkisini modellemeyi amaçlar. Bu modelleme süreci, veri setindeki değişkenlikleri anlama ve gelecekteki değerleri tahmin etme amacı taşır. Regresyon yöntemleri, veri analizi ve öngöründe güçlü bir araç olarak öne çıkar.

Bu tez çalışmasında, spor performansı ile ilişkilendirilen değerleri tahmin etme amacıyla makine öğrenimi regresyon algoritmalarının kullanılacağı bir perspektif sunmayı amaçlamaktadır. YDT gibi spesifik veri setlerinin analizi, sporcuların dengesini ve motor koordinasyonunu değerlendirmede önemli bir rol oynamaktadır. Bu veriler, sporcuların performans potansiyelini anlamada ve potansiyel risk faktörlerini belirlemede kritik bir rol oynamaktadır. Makine öğrenimi regresyon algoritmaları, bu tür veri setlerindeki karmaşık ilişkileri analiz etme ve gelecekteki performans tahmin etme yetenekleriyle dikkat çekmektedir.

Bu çalışma, farklı regresyon yöntemlerinin YDT verilerinin tahmini için kullanımını inceleyerek, sporcuların performansını optimize etmek ve yaralanma riskini azaltmak için değerli içgörüler sunmayı amaçlamaktadır. Bu şekilde, veri odaklı bir yaklaşım kullanarak spor alanında daha etkili kararlar alınmasına katkı sağlanması hedeflenmektedir. Ağaç tabanlı regresyon modellerinin tercih edilmesinin altında yatan sebep, tez çalışmasının temel amacıyla ve veri setinin özellikleriyle uyumlu bir analitik yaklaşım sağlamış olmasıdır. Bu modeller, geniş bir uygulama yelpazesi ve esneklik sunan özelliklere sahip olmaları nedeniyle tercih edilmiştir. Veri setindeki özelliklerin doğrusal ve doğrusal olmayan ilişkiler içermeye potansiyeline sahip olduğu göz önüne alındığında, ağaç tabanlı modellerin bu esneklikten yararlanabileceği düşünülmüştür. Ek olarak, veri setinde aykırı değerlerin varlığı, ağaç tabanlı

modellerin bu aykırı deęerlere karřı daha dayanıklı olduęu bilgisi nedeniyle göz önünde bulundurulmuřtur.

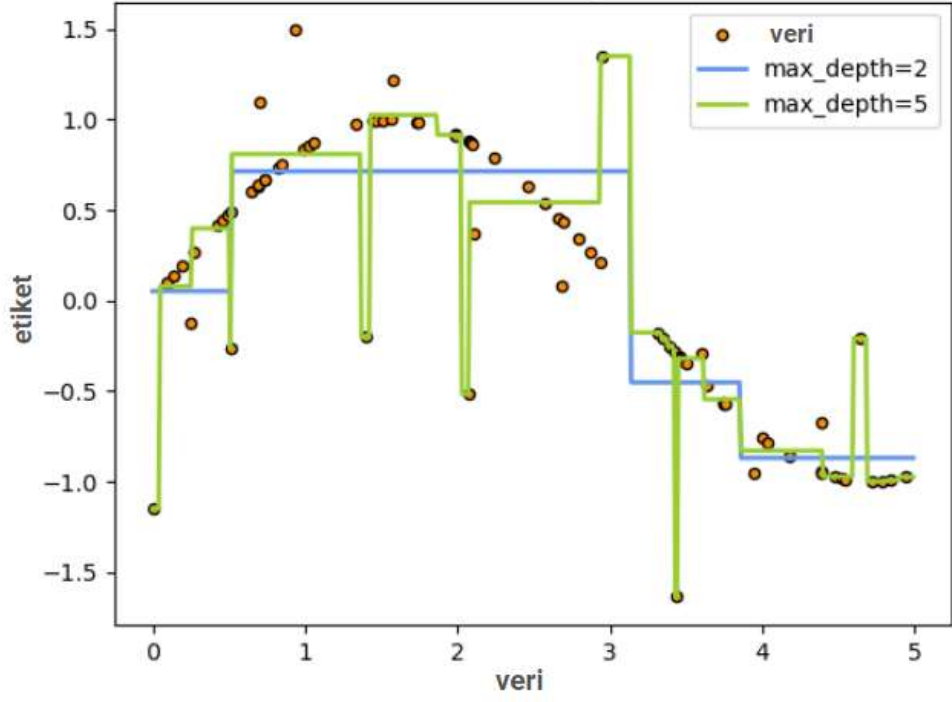
Bu modeller, her bir ağacın tahminlerini birleřtirerek topluluk (ensemble) öğrenimi sağladıęı için, daha iyi genel bir performans elde etme potansiyelini beraberinde getirmiřtir. Bu nedenle, tezde elde edilecek sonuçların yüksek tahmin doęruluęuna sahip olmanın yanı sıra genellemeye de uygun olacaęına inanılmaktadır.

### **3.2.1. Karar Ağaçları (Decision Trees) Regresyon**

Karar Ağaçları Regresyon, makine öğrenimi alanında yaygın olarak kullanılan bir regresyon yöntemidir. Bu yöntem, veri setindeki özellikleri kullanarak ağaç yapısı şeklinde karar verme kuralları oluşturur. Her bir iç düęümde bir özellik seçilir ve bu özellięe göre veri kümesi bölünür. Her dalda bir sonuç elde edilene kadar bu bölünmeler devam eder. Karar Ağaçları Regresyon, bağımsız deęişkenlerin sürekli deęerler alabileceęi durumlar için kullanılır. Her bir yaprak düęümde ise ortalama veya ağırlıklı ortalama deęeri hesaplanarak tahminde bulunulur. Bu sayede, veri setindeki karmařık iliřkileri anlamak ve tahmin yapmak mümkün hale gelir [60].

Karar Ağaçları Regresyon algoritması, özellikle doğrusal olmayan iliřkileri yakalamak ve veri setindeki içsel yapılara anlam kazandırmak için etkili bir araçtır. Ayrıca, bu yöntem sayesinde özelliklerin önemi de deęerlendirilebilir ve karar ağacının görselleřtirilmesi ile sonuçlar daha anlaşılır bir şekilde sunulabilir. Karar Ağaçları Regresyonun avantajları arasında kolay anlaşılır sonuçlar üretebilmesi ve veri ön işleme ihtiyacının düşük olması yer alır. Ancak, aşırı dallanma durumunda overfitting riski bulunmaktadır. Bu nedenle ağacın uygun derinlikte kesilmesi önemlidir. Bu duruma Şekil 3.5'teki görseldeki gibi bir örnek verebiliriz.

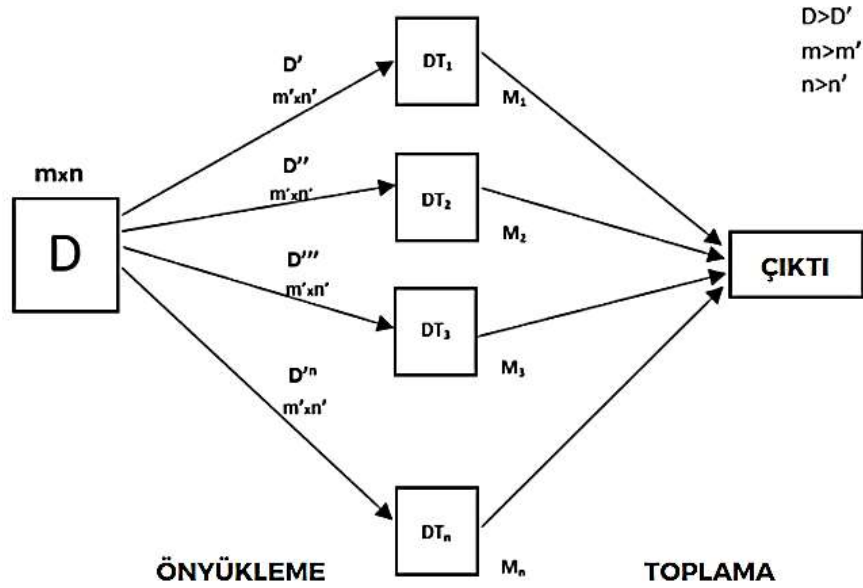




Şekil 3.5. Karar ağaçları regresyon algoritmasının örnek çıktısı [61].

### 3.2.2 Rastgele Orman (Random Forest) Regresyon

Rastgele Orman Regresyon, makine öğrenimi alanında yaygın olarak kullanılan ve yüksek performanslı tahminler yapabilen bir regresyon yöntemidir. Temel olarak birçok karar ağacının bir araya gelmesiyle oluşan bir topluluk (ensemble) modelidir [62]. Şekil 3.6’da bu yapıya bir örnek verilmiştir. Bu algoritma hem regresyon hem de sınıflandırma problemleri için etkili bir şekilde kullanılabilir.



Şekil 3.6. Rastgele orman regresyon algoritmasının çalışma prensibi [63].

Rastgele Orman Regresyon, veri kümesini parçalara ayırarak her parçada ayrı bir karar ağacı oluşturur. Bu karar ağaçları, veri noktalarını farklı özelliklere göre böler ve her bir özelliğin çıktı üzerindeki etkisini değerlendirir. Ardından, bu ağaçlar bir araya getirilerek tahminler üretilir. Her bir ağacın tahminleri ağırlıklandırılarak bir araya getirilir ve topluluk (ensemble) modeli oluşturulur. Bu yöntem sayesinde aşırı uyum riski azalırken, modelin genel performansı artırılır.

Rastgele Orman Regresyon avantajları arasında veriye ait özelliklerin önemini belirleme yeteneği, aykırı değerlere karşı dayanıklılığı ve yüksek boyutlu veri setlerinde iyi performans gösterme yeteneği yer alır. Ayrıca, modelin genel performansını artırmak ve varyansı düşürmek için ağaç sayısı, maksimum derinlik ve özellik seçimi gibi parametreler ayarlanabilir.

### 3.2.3. Adaptif Artırma (Adaptive Boosting) Regresyon

Adaptif Artırma (AdaBoost) Regresyon, tahmin yapmak ve regresyon problemlerini çözmek için kullanılan bir makine öğrenimi algoritmasıdır. Temel amacı, zayıf

tahmincilerin (örneğin, düşük derinlikli karar ağaçları) bir araya getirilerek daha güçlü bir tahminci oluşturulmasıdır.

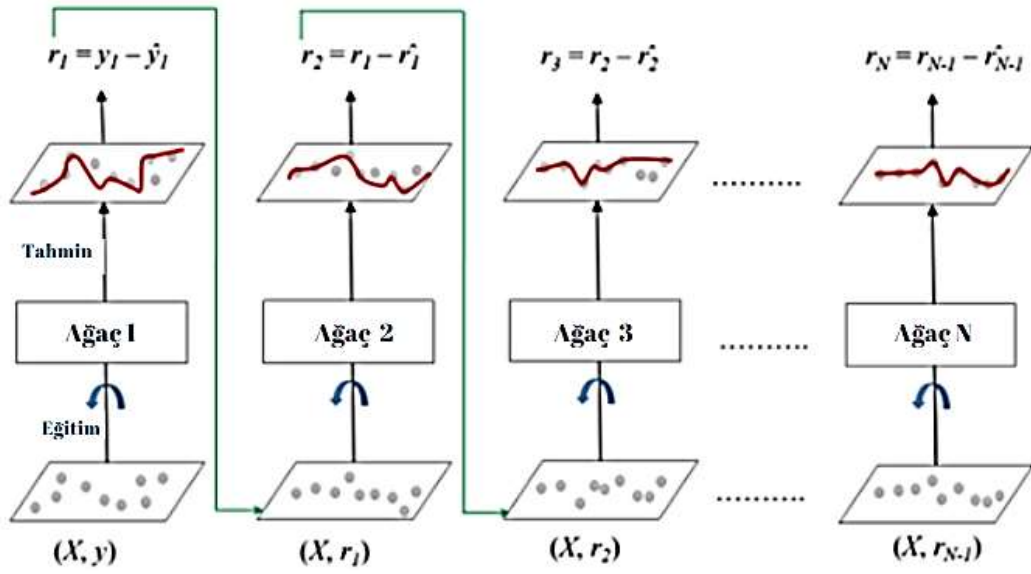
Algoritma, veri setindeki her örnek için başlangıçta eşit ağırlıklarla başlar. Ardından, bir zayıf tahminci oluşturulur ve bu tahminci hatalı tahmin edilen örneklerin ağırlığını artıracak şekilde ayarlanır. Bu, hatalı tahmin edilen örneklerin daha fazla dikkate alınmasını sağlar ve bu hataların düzeltilmesi hedeflenir. Daha sonra, bir sonraki zayıf tahminci oluşturulur ve bu tahminci de hatalı tahmin edilen örneklerin ağırlığını artıracak şekilde güncellenir. Bu adım, önceki tahmincilerin hatalarına odaklanarak bir dizi tahminci oluşturulmasını sağlar. Son tahmin, her bir tahmincinin çıktısının ağırlığına göre hesaplanır. Hatalı tahmin edilen örneklerin ağırlığı daha yüksek olduğu için son tahmin, bu örneklerin daha iyi tahmin edilmesine odaklanır [60].

AdaBoost Regresyonu, bu şekilde zayıf tahmincileri birleştirerek daha güçlü ve doğru bir tahminci oluşturur. Bu algoritma, veri setindeki karmaşıklığı ele alır ve düşük doğruluğa sahip tahmincilerin birleştirilmesiyle yüksek doğruluklu tahminler elde edilmesini sağlar.

#### **3.2.4. Gradyan Artırma (Gradient Boosting) Regresyon**

Gradyan Artırma Regresyon, makine öğrenimi alanında yaygın olarak kullanılan bir algoritmadır ve özellikle regresyon problemlerinde etkili sonuçlar verir. Bu algoritma, birçok zayıf tahminleyici modelini (genellikle karar ağaçları) birleştirerek güçlü bir tahminleyici modeli oluşturmayı amaçlar [64]. Her bir zayıf tahminleyici, önceki tahmincilerin hatalarını düzeltmeye çalışarak oluşturulur.

Gradyan Artırma, adından da anlaşılacağı üzere, bir gradyan iniş optimizasyon algoritması kullanır. Her bir tahminci, gerçek çıktı değerleri ile tahminler arasındaki farkı minimize etmeye çalışarak eğitilir. Bu şekilde her adımda model, önceki adımların hatalarını telafi eder ve daha güçlü bir tahminleyici oluşturulur. Şekil 3.7'de bu duruma örnek bir görsel verilmiştir.



Şekil 3.7. Gradyan artırma regresyon algoritmasının çalışma prensibi [65].

Algoritmanın en önemli parametresi, tahmincilerin sayısı ve derinliğidir. Bu parametreler, aşırı öğrenme ve alt öğrenme arasında bir denge kurmayı gerektirir.

### 3.2.5. Aşırı Gradyan Artırma (Extreme Gradient Boosting) Regresyon

Aşırı Gradyan Artırma (XGBoost) Regresyon, makine öğrenimi alanında güçlü bir regresyon algoritması olarak öne çıkan ve geniş bir uygulama yelpazesi sunan bir yöntemdir. Temel amacı, zayıf tahmincileri birleştirerek daha güçlü ve yüksek performanslı bir tahminci oluşturmaktır [66].

Algoritma, gradyan iniş yöntemine dayalıdır ve her adımda hata fonksiyonunun gradyanına odaklanarak tahmincileri günceller. İlk adımda, veri setindeki her örnek için başlangıçta eşit ağırlıklarla bir tahminci oluşturulur. Sonrasında, her adımda yeni bir tahminci eklenir ve bu tahminci önceki tahmincilerin hatalarını düzeltmeye çalışır.

XGBoost, her bir tahminciyi ağaç yapısı olarak temsil eder. Bu ağaçlar, veri setindeki özelliklerin ve hedef değerlerin ilişkilerini yakalamak için tasarlanır. Algoritma, bu ağaçları birleştirerek daha güçlü ve genelleyici bir model oluşturur.

Ayrıca, XGBoost'un aşırı öğrenmeye karşı dirençli olma yeteneği ve hiperparametrelerin optimize edilmesi için çeşitli araçlar sunması da önemli avantajları arasında yer alır. Bu sayede, modelin hem karmaşıklığı hem de performansı optimal bir şekilde dengelemesi sağlanır.

XGBoost, büyük veri setleri üzerinde yüksek tahmin doğruluğu elde etme yeteneği ile öne çıkar. Çeşitli hedef fonksiyonlarını ve regresyon problemlerini ele alabilir. Bu nedenle, araştırmacılar ve veri bilimciler için güvenilir bir araç olarak tercih edilir.

Sonuç olarak, XGBoost regresyon algoritması, zayıf tahminicileri bir araya getirerek güçlü ve doğru tahminler yapmayı sağlayan etkili bir yöntemdir.

### **3.2.6. Stokastik Gradyan İniş (Stochastic Gradient Descent) Regresyon**

Stokastik Gradyan İniş (SGD) Regresyon, makine öğreniminde sıkça kullanılan bir regresyon algoritmasıdır. Temel amacı, veri kümesindeki ilişkileri anlamak ve öngörülerde bulunmak için model oluşturmaktır [67]. SGD Regresyon, özellikle büyük veri kümeleri üzerinde etkili bir şekilde çalışabilen ve hızlı hesaplamalarla sonuç elde edebilen bir algoritmadır.

Algoritma, her adımda bir örnek veya alt küme üzerinden gradient iniş yöntemini uygular. Bu şekilde, modeli güncellerken örnekleri veya alt kümeleri rastgele seçerek hızlı bir şekilde optimize etmeye çalışır. Gradient iniş, öngörülerini gerçek değerlere yaklaştırmak için gradyan (eğim) hesaplamalarını kullanır ve bu şekilde tahmin hatalarını azaltmayı amaçlar.

## BÖLÜM 4

### Y-DENGE TESTİ VERİLERİNİN MAKİNE ÖĞRENİMİ ALGORİTMALARI İLE ANALİZİ

Günümüzde veri analizi, bilgi çıkarmada ve karar verme süreçlerinde hayati bir rol oynamaktadır. Bu bağlamda, YDT verileri gibi kapsamlı ve karmaşık veri setlerinin etkili bir şekilde analiz edilmesi, değerli bilgiler elde etmek için kritik bir öneme sahiptir. Bu tez çalışması, YDT verilerinin makine öğrenimi algoritmalarıyla analiz edilmesini ele almaktadır. Veri analizi süreci, veri temizleme, normalizasyon, algoritma seçimi, performans değerlendirmesi ve sonuçlar aşamalarını içermektedir. Çalışmanın amacı, YDT verilerinin içerdiği bilgileri daha iyi anlamak, desenleri ortaya çıkarmak ve gelecekteki kararları desteklemek için uygun makine öğrenimi algoritmalarını tanımlamak ve değerlendirmektir. Çalışmanın ilerleyen bölümlerinde, veri analizi sürecinin her aşaması ayrıntılı olarak ele alınacak ve YDT verileri üzerinde uygulanacak makine öğrenimi algoritmalarının performansı incelenecektir.

#### 4.1. VERİ ÖN İŞLEME

Veri analizi ve makine öğrenimi projeleri, karmaşık ve geniş veri kümelerini içerirken, bu verilerin doğru ve güvenilir bir şekilde işlenmesi gerekmektedir. Veri setlerinin çeşitliliği, eksik verilerin varlığı ve özelliklerin farklı ölçeklerde olması gibi zorluklar, verinin doğru bir şekilde anlaşılmasını ve kullanılmasını engelleyebilir. İşte bu noktada veri ön işleme adımı devreye girer. Veri ön işleme, verinin temizlenmesi, düzenlenmesi ve hazırlanması amacıyla gerçekleştirilen önemli bir adımdır. Eksik verilerin yönetilmesi, veri normalizasyonu ve kategorik verilerin sayısal değerlere dönüştürülmesi gibi işlemler, veri setinin kalitesini artırarak model performansını iyileştirmeyi amaçlar. Bu girişim, veri analizi ve makine öğrenimi projelerinde sağlıklı sonuçlar elde etmek için temel bir adımdır.

Çalışmada, veri ön işleme sürecini daha iyi ele almak ve bu adımları daha etkili bir şekilde açıklamak amacıyla Python programlama dili ve Jupyter Notebook ortamı kullanılmıştır.

#### 4.1.1. Veri Setini Yükleme ve Eksik Verilerin İşlenmesi

Analiz aşamasında Şekil 4.1’deki gibi kullanacağımız kütüphanelerin import edilmesi ile analiz süreci başlar.

```
# Kütüphanelerin import edilmesi

import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

from lazypredict.Supervised import LazyClassifier, LazyRegressor
from sklearn.preprocessing import MinMaxScaler

from sklearn.ensemble import RandomForestRegressor
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import GradientBoostingRegressor
from xgboost import XGBRegressor

from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score
```

Şekil 4.1. Gerekli kütüphanelerin import edilmesi.

CSV uzantılı veri setini yükledikten sonra veri setindeki verilerin tiplerinin kontrolü ve eksik verilerin kontrolü sağlanmıştır. Şekil 4.2’de veri setinin içeriğini görmekteyiz.

Cinsiyet	Branş	Yaş	BMI	Antrenman Yaşı	SAGANTR%	SOLANT%	ANT%ORT	SAGPM%	SOLPM%	PM%ORT	SAGPL%	SOLPL%	PL%ORT	TOPLAMORT%	
0	K	TENİS	6.00	13.44	1.00	51.79	54.36	53.06	85.13	87.18	86.15	77.95	91.78	84.87	74.70
1	K	TENİS	7.00	17.53	1.00	50.72	49.76	50.24	65.22	76.33	70.77	55.56	71.98	63.77	61.59
2	K	TENİS	7.00	13.64	2.00	44.00	44.13	44.37	95.31	89.67	92.49	71.30	70.42	70.89	89.25
3	K	TENİS	7.00	16.00	2.00	60.83	55.56	58.20	127.51	119.05	123.28	107.94	81.48	94.71	92.06
4	K	TENİS	7.00	18.93	1.00	51.21	52.17	51.69	105.80	77.29	91.55	84.20	60.87	77.54	73.59
212	K	BASKETBOL	15.00	19.66	5.00	57.28	49.19	53.24	76.38	75.40	75.89	75.08	87.38	81.23	70.12
213	E	BASKETBOL	14.00	18.62	6.00	58.00	40.00	59.00	100.33	89.00	94.67	77.33	75.00	76.17	76.61
214	E	BASKETBOL	15.00	21.03	9.00	57.33	53.67	55.50	91.67	91.67	91.67	73.33	75.67	74.50	73.88
215	E	BASKETBOL	15.00	19.59	4.00	65.99	64.98	65.49	92.26	83.84	89.05	81.14	84.18	82.68	78.73
216	E	BASKETBOL	15.00	28.65	6.00	54.41	49.43	51.92	86.59	87.74	87.16	82.38	80.46	81.42	73.50

217 rows x 15 columns

Şekil 4.2. Veri setinin örnek içeriği.

Eksik veriler, veri analizi süreçlerinde yaygın bir sorundur. Eksik verileri tanımlamak, bu verileri doldurmak veya çıkarmak önemlidir. Bu adım, veri bütünlüğünü ve güvenilirliğini artırmaya yönelik atılan temel bir adımdır [68]. Kullanılan veri setinde eksik olan veriler çıkartılarak devam edilmiştir.

*df.dropna(inplace=True)*

*df.info()*

Şekil 4.3'te görüldüğü üzere, veri setinde cinsiyet ve branş gibi değerler kategorik veri olarak yer almaktadır. Kategorik verilerin dönüştürülmesi, veri analizi ve makine öğrenimi süreçlerinde büyük bir öneme sahiptir. Genellikle sembolik veya metinsel değerlerle ifade edilen kategorik veriler, pek çok algoritma tarafından doğrudan işlenmeye uygun değildir. Bu nedenle, kategorik verilerin sayısal değerlere dönüştürülmesi, algoritmaların bu verileri daha etkin bir şekilde işlemesine imkan tanır. Bu dönüşüm aynı zamanda istatistiksel analizler ve özellik mühendisliği süreçleri için de kullanılabilir hale gelir.



#	Column	Non-Null Count	Dtype
0	Cinsiyet	217 non-null	object
1	Branş	217 non-null	object
2	Yaş	217 non-null	float64
3	Boy	217 non-null	float64
4	Kilo	217 non-null	float64
5	BMI	217 non-null	float64
6	Antrenman Yaşı	217 non-null	float64
7	Baskın Bacak	217 non-null	int64
8	SAĞANTR%	217 non-null	float64
9	SOLANT%	217 non-null	float64
10	ANT%ORT	217 non-null	float64
11	SAĞPM%	217 non-null	float64
12	SOLPM%	217 non-null	float64
13	PM%ORT	217 non-null	float64
14	SAĞPL%	217 non-null	float64
15	SOLPL%	217 non-null	float64
16	PL%ORT	217 non-null	float64
17	TOPLAMORT%	217 non-null	float64

Şekil 4.3. Kategorik verilerin tespit edilmesi.

#### 4.1.2 Normalizasyon İşlemi

Normalizasyon, farklı özelliklerin farklı ölçeklerde olabileceği durumları dengelemek ve modelin daha istikrarlı bir şekilde çalışmasını sağlamak amacıyla yapılır. Bu adım, özellikle öklidyen uzaklık gibi metriklerle dayalı algoritmaların veri üzerinde dengeli bir şekilde çalışmasını sağlamak için önemlidir [69]. Min-Max Ölçeklendirme veya Z-Skor Normalizasyon gibi yöntemler kullanılarak veri, belirli bir aralığa veya dağılıma getirilir. Bu adım, veri madenciliği ve makine öğrenimi projelerinde yaygın bir şekilde kullanılır [70].

Min-Max Ölçeklendirme yöntemi, verinin minimum ve maksimum değerleri arasında belirli bir aralığa dönüştürülmesini sağlar. Bu yöntemde her bir özellik değeri, verinin minimum ve maksimum değerleri arasında kalan bir değere yeniden ölçeklenir. Bu sayede tüm özellikler aynı aralığa sahip olur ve algoritmaların farklı ölçeklerdeki özellikleri eşit şekilde değerlendirmesi sağlanır [69]. Eşitlik 4.1'de min-max normalizasyonun formülü verilmiştir.

$$X_{sc} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4.1)$$

Kod içerisinde Sklearn kütüphanesinden çağırarak kullanabiliriz.

```
from sklearn.preprocessing import MinMaxScaler  
scaler = MinMaxScaler()
```

Z-Skor Normalizasyonu, özellik değerlerini ortalama değere göre merkezlemek ve standart sapma ile bölerek dönüştürmek amacıyla kullanılır. Her bir özellik değeri, ortalama değerden çıkarılır ve standart sapmaya bölünerek dönüştürülür. Bu yöntem, özelliklerin dağılımını daha benzer bir hale getirir ve uzaklık tabanlı algoritmaların veriyi daha dengeli bir şekilde işlemesine yardımcı olur [68].

Veri noktalarını Z-skorlara dönüştürmek için kullanılan formül Eşitlik 4.2’de gösterilmiştir.

$$z = \frac{x - \mu}{\sigma} \quad (4.2)$$

- $z$  veri noktasının Z-skorunu temsil eder.
- $x$  orijinal veri noktasını temsil eder.
- $\mu$  veri setinin ortalama değerini temsil eder.
- $\sigma$  veri setinin standart sapmasını temsil eder.

Min-Max Ölçeklendirme yöntemi veri setinin dağılımını bozmadan ve özelliklerin anlamını değiştirmeden uygulanabilir olması nedeniyle tercih edilebilir. Bu sayede model performansını artırarak daha güvenilir sonuçlar elde etmeyi amaçlar. Çalışmada Min-Max Ölçeklendirme yöntemini kullanarak veri setini ön işleme aşamasında ölçeklendirme yapılması, algoritmaların daha etkili ve istikrarlı bir şekilde çalışmasını hedeflerken, özelliklerin orijinal anlamını ve dağılımını bozmadan veri setini hazırlamayı amaçlamaktadır.

### 4.1.3. Kategorik Verilerin İşlenmesi

Kategorik değişkenler, sayısal olmayan ve genellikle sınıf veya kategori türünden olan veri noktalarını ifade eder. Bu tür değişkenler, makine öğrenimi algoritmalarına doğrudan uygulanamaz, çünkü algoritmalar sayısal veri gerektirir. Bu nedenle, kategorik değişkenleri sayısal verilere dönüştürmek için farklı yöntemler kullanılır. İki yaygın yöntem One-Hot kodlama ve Label kodlama olarak adlandırılır.

One-hot kodlama, kategorik değişkenleri ikili (0 veya 1) değerlere dönüştürerek sayısal verilere dönüştürme yöntemidir. Bu yöntem, model performansını artırmak ve ağaç tabanlı modellerin kategorik verileri daha iyi işlemesini sağlamak amacıyla kullanılır [71].

One-Hot kodlama, her kategoriye ait yeni bir özellik oluşturarak kategorik değişkenleri sayısal hale getirir. Her kategori özelliği 0 veya 1 ile temsil edilir, bu da veri noktalarının sadece bir kategoriye ait olduğunu gösterir. Kullanılan veri setinde Cinsiyet ve Branş kategorik verilerdir. Bu veriler one-hot kodlama uygulanarak sayısal verilere dönüştürülmüştür. Cinsiyette 1'ler kadın sporcuları 0'lar ise erkek sporcuları temsil etmektedir. Çizelge 4.1'de veri setine one-hot kodlama işlemi uygulandıktan sonraki hali gösterilmiştir.

Çizelge 4.1. One-Hot kodlama uygulanmış veriler.

Cinsiyet	BASK.	CİMN.	FUTB.	HENT.	TEKV.	TENİS	VOLE.	YÜZME
1	0	0	0	0	0	1	0	0
1	0	1	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	0	1	0
0	0	0	0	1	0	0	0	0
0	0	0	1	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1

One-Hot kodlama kullanmanın avantajları arasında algoritmaların kategorik değişkenleri daha iyi anlaması ve kategorik değişkenlerin aralarındaki ilişkileri yansıtması yer almaktadır.

Diğer bir yöntem olan Label kodlama, kategorik değişkenleri sıralı bir şekilde sayılarla temsil etmeye dayanır. Her kategoriye benzersiz bir sayı atanır. Ancak bu yöntemin dikkat edilmesi gereken önemli bir noktası, algoritmaların sıralama bilgisini yanlış yorumlayabileceğidir. Yani algoritma, kategorik değişkenin sayısal temsilini sıralama gibi algılayabilir, bu nedenle bu yöntem sadece bazı modellerde tercih edilir. Label kodlama kullanmanın avantajları arasında basit ve hızlı bir yöntem olması ile bellek kullanımında daha az yer gerektirmesi yer almaktadır.

## **4.2. MODEL SEÇİMİ**

Bu tez çalışması, makine öğrenimi algoritmalarının genel başarısını değerlendirmek amacıyla çeşitli regresyon algoritmalarını içermektedir. Bununla birlikte, algoritmaların karşılaştırmalı analizi sadece geleneksel regresyon algoritmaları ile sınırlı değildir. Bu çalışmanın değerlendirme sürecine “Tembel Tahmin Regresyon (Lazy Predict Regressor)” adlı etkili bir araç da dahil edilmiştir. Bu araç, özellikle geniş bir veri kümesi içerisinde en iyi performansı gösteren regresyon algoritmasını belirlemek amacıyla kullanılan bir otomatik model seçimi aracıdır [72].

Tembel Tahmin Regresyon, çeşitli regresyon algoritmalarının genel performansını tahmin etmek için kullanılan bir araçtır. Bu araç, veri kümesindeki özellikleri analiz eder, algoritmaları otomatik olarak seçer ve her bir algoritmanın tahmin performansını önceden tahmin eder. Bu sayede, veri bilimi uygulayıcılarına hangi algoritmaların daha iyi performans gösterebileceği konusunda hızlı bir bakış sunar. Tembel Tahmin Regresyon'dan elde edilen sonuçlara dayanarak toplu (ensemble) modeller veri setimizde daha yüksek sonuçlar vermiştir. Bu yüzden bu modelleri ayrı ayrı ele alarak daha detaylı bir analize gidilmiştir.

## BÖLÜM 5

### DENEYSEL ÇALIŞMALAR

Makine öğrenimi algoritmalarının performans analizini gerçekleştirmek amacıyla çalışmada Jupyter Notebook ortamı kullanılmıştır. Bu bölümde, çeşitli regresyon algoritmalarının karşılaştırmalı analizi yapılmıştır. Bu analiz, sağ antr, sol antr, sağ pm, sol pm, sağ pl ve sol pl değerlerinin tahminlenmesi amacıyla 6 farklı regresyon modelinin performansını incelemeyi hedeflemiştir.

Bağımlı değişkenlerin farklı kombinasyonları kullanılarak toplamda 6 ayrı regresyon modeli oluşturulmuştur. Bu modeller, her bir bağımlı değişken için en uygun tahminleme algoritmasını belirlemek amacıyla değerlendirilmiştir. Veri kümesi, gelişmiş regresyon algoritmaları olan Random Forest Regressor, Decision Tree Regressor, AdaBoost Regressor, Gradient Boosting Regressor, XGBoost Regressor ve SGD Regressor algoritmaları kullanılarak tahminlenmiştir.

Çalışmada, kullanılan algoritmaların parametreleri varsayılan (default) değerlerine ayarlanmıştır. Random Forest Regressor'da ağaç sayısı 100, maksimum derinlik otomatik olarak belirlenmiş ve bölünme kriteri MSE olarak kullanılmıştır. Decision Tree Regressor'da da maksimum derinlik otomatik olarak belirlenirken, bölünme kriteri yine MSE olarak kullanılmıştır. AdaBoost Regressor'da ise ağaç sayısı 50, öğrenme oranı 1.0 ve tahminci tipi Decision Tree olarak ayarlanmıştır. Gradient Boosting Regressor'da ağaç sayısı 100, öğrenme oranı 0.1 ve bölünme kriteri MSE olarak belirlenmiştir. XGBoost Regressor'da varsayılan olarak maksimum derinlik 6, öğrenme oranı 0.3 ve ağaç sayısı 100 kullanılmıştır. Son olarak, SGD Regressor'da ise varsayılan olarak öğrenme oranı 0.01, yineleme sayısı 1000 ve kayıp fonksiyonu "squared\_loss" olarak ayarlanmıştır. Bu parametreler, her bir algoritmanın varsayılan yapılandırmalarını yansıtmaktadır.

Her bir modelin performansı, hata metrikleri olan MSE, R2 ve RMSE kullanılarak değerlendirilmiştir. Bu metrikler, modellerin tahmin yeteneğini ve tahmin sonuçlarının gerçek değerlere ne kadar yakın olduğunu ölçmek için kullanılmıştır.

Sonuçlar, her bir modelin sağ ant, sol ant, sağ pm, sol pm, sağ pl ve sol pl değerlerini tahminlemedeki performansını ortaya koymaktadır. Analiz sonuçlarına göre, her bir bağımlı değişken için en iyi tahminleme performansını gösteren model belirlenmiştir. Elde edilen sonuçlar, bu çalışmanın hedeflerine ulaşılmasında makine öğrenimi algoritmalarının etkisini ve performansını açıkça göstermektedir.

Bu analiz, regresyon algoritmalarının çeşitli bağımlı değişken kombinasyonları üzerindeki etkisini ayrıntılı bir şekilde incelemekte ve her bir algoritmanın performansını objektif bir şekilde değerlendirmektedir. Bu sonuçlar, gelecekte benzer veri kümesi üzerinde tahminleme yaparken hangi algoritmanın tercih edilmesi gerektiği konusunda rehberlik edici olabilir.

Eşitlik 5.1'deki MSE, bir regresyon modelinin performansını değerlendirmek için kullanılan bir metriktir. Bu metrik, modelin tahmin ettiği değerler ile gerçek değerler arasındaki farkları hesaplar, bu farkların karesini alır ve bu kare farkların ortalamasını bulur. MSE'nin düşük olması, modelin tahminlerinin gerçek değerlere yakın olduğunu ve iyi bir uyum sağlandığını gösterir. Ancak MSE'nin orijinal verilerin birim karesinde ifade edilmesi nedeniyle veri ölçeğine duyarlıdır.

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (5.1)$$

Eşitlik 5.2'deki R2 (R-kare) metriği ise, regresyon modelinin bağımsız değişkenler kullanılarak bağımlı değişkenin varyansının ne kadarını açıkladığını ölçer. R2 değeri, 0 ile 1 arasında değer alır. R2'nin yüksek olması, modelin bağımlı değişkendeki varyansın büyük bir kısmını açıkladığını ve iyi bir uyum sağlandığını gösterir. Ancak R2 metriği, aşırı uyum sorununa yol açabilir; yani fazla karmaşık modellerde R2 değeri yüksek çıkabilir ancak bu modelin genelleme yeteneğini düşürebilir.

$$R^2 = c^T R_{xx}^{-1} c \quad (5.2)$$

Eşitlik 5.3'teki Kök RMSE, MSE'nin kökünü alarak elde edilen bir metriktir. RMSE, MSE'nin ölçeklendirilmiş bir versiyonu olarak düşünülebilir. Bu metrik, tahmin edilen değerler ile gerçek değerler arasındaki farkların ortalama büyüklüğünü hesaplar ve bu farkların tahminlerin ne kadar hassas olduğunu gösterir. RMSE de düşük olduğunda, modelin tahminleri gerçek değerlere yakın bir şekilde dağıldığını ifade eder.

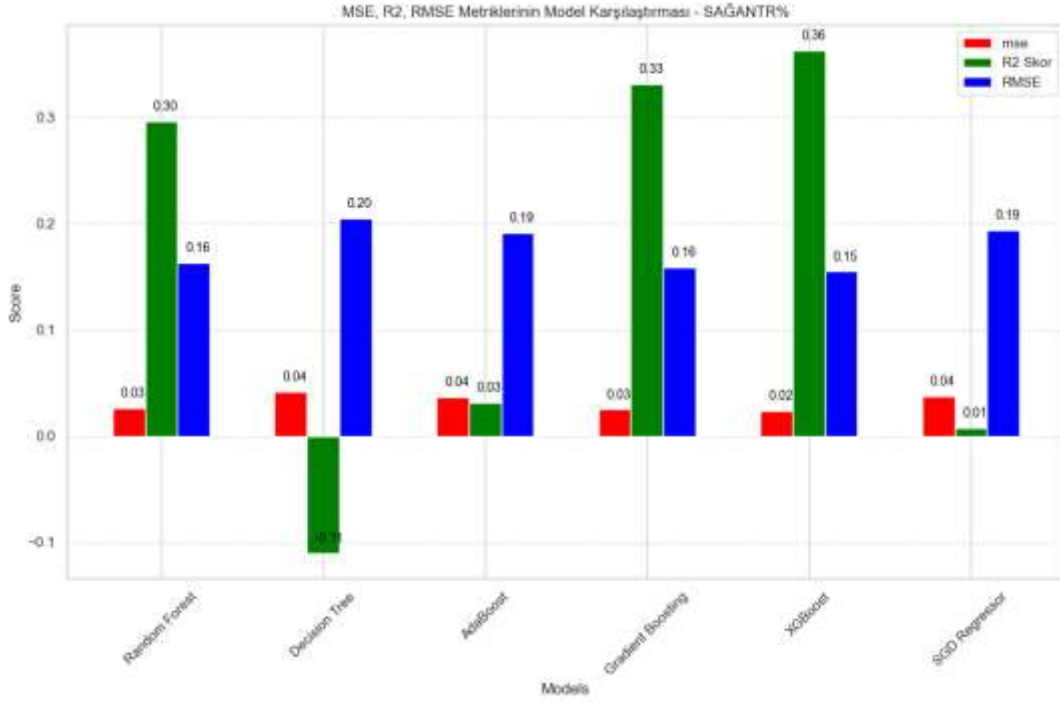
$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (5.3)$$

Bu metrikler, regresyon modellerinin performansını ölçmek ve karşılaştırmak için önemlidir. Veri analizi ve makine öğrenimi çalışmalarında, bu metriklerin dikkatli bir şekilde yorumlanması ve kullanılması, model seçiminde ve iyileştirmelerde rehberlik sağlar.

Çizelge 5.1'de ve Şekil 5.1'de SAĞANTR% tahmini için algoritmaların performansı gösterilmiştir. En iyi sonucu 0,02 MSE, 0,36 R2 ve 0,15 RMSE değeri ile XGBoost algoritması vermiştir.

Çizelge 5.1. SAĞANTR% değeri tahmininde regresyon algoritmaları sonuçları.

Değişkenler	Hedef YDT değeri	Metrik	Makine Öğrenimi Regresyon Algoritmaları					
			Decision Tree	Random Forest	Ada Boost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş	SAĞANTR %	MSE	0,04	0,03	0,04	0,03	<b>0,02</b>	0,04
		R2	-0,11	0,30	0,03	0,33	<b>0,36</b>	0,01
		RMSE	0,20	0,16	0,19	0,16	<b>0,15</b>	0,19



Şekil 5.1. SAĞANTR% değeri tahmininde model performansları.

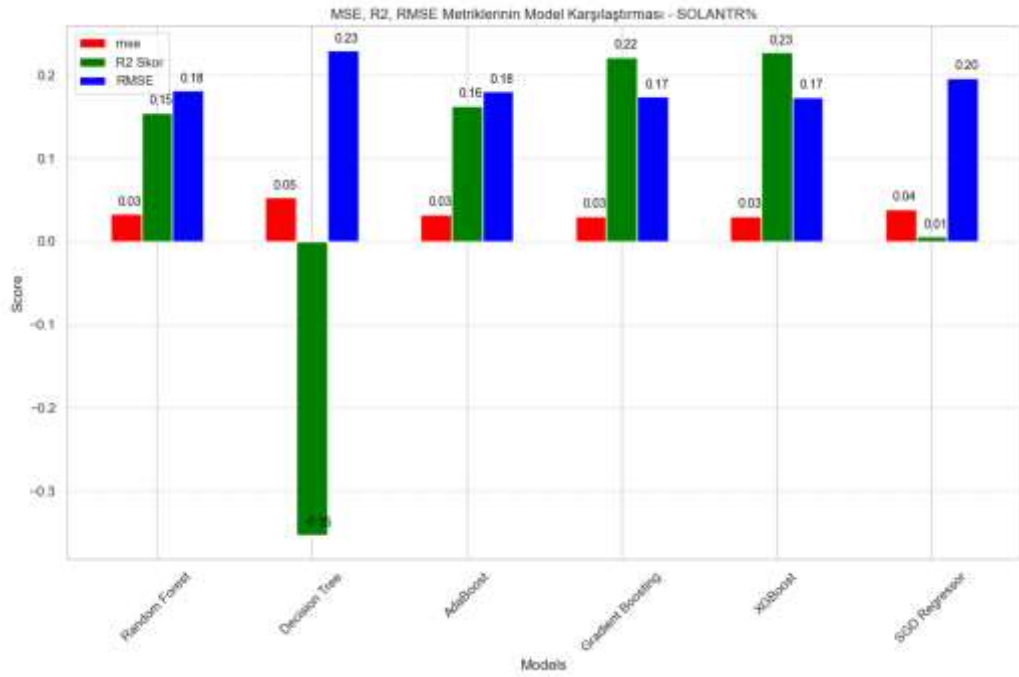
Random Forest, Gradient Boosting, XGBoost algoritmalarının burada diğer algoritmalara göre daha iyi sonuçlar verdiğini gözlemlemekteyiz. Burada görülen bir diğer durum ise Decision Tree modelinin R2 skorunun negatif değerlere sahip olmasıdır. Bu durum, modelin verilere yetersiz uyum sağladığını ve hatta bazı durumlarda verileri rastgele tahmin ettiğini gösterir. Negatif R2 skoru, modelin verilerdeki örüntüleri yakalayamadığını ve tahminlerinin gerçek değerlere kıyasla ciddi bir sapma gösterdiğini gösterir. Bu sonuç, modelin yapılandırılmasında veya veri özelliklerinin seçiminde olası bir sorun olduğunu düşündürmektedir. Karar ağaçları algoritmasının yapısı, veri setinin karmaşıklığı veya kullanılan hiperparametreler gibi etkenler negatif R2 sonucuna katkıda bulunmuş olabilir. Bu sonuçların dikkate alınarak modelin iyileştirilmesi veya farklı bir modelin seçilmesi gerekebilir.

Çizelge 5.2’de ve Şekil 5.2’de SOLANTR% tahmini için algoritmaların performansı gösterilmiştir. En iyi sonucu 0,03 MSE, 0,23 R2 ve 0,17 RMSE değeri ile XGBoost algoritması vermiştir.



Çizelge 5.2. SOLANTR% değeri tahmininde regresyon algoritmaları sonuçları.

			Makine Öğrenimi Regresyon Algoritmaları					
Değişkenler	Hedef YDT değeri	Metrik	Decision Tree	Random Forest	AdaBoost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş	SOL ANTR %	MSE	0,05	0,03	0,03	0,03	<b>0,03</b>	0,04
		R2	-0,35	0,15	0,16	0,22	<b>0,23</b>	0,01
		RMSE	0,23	0,18	0,18	0,17	<b>0,17</b>	0,20



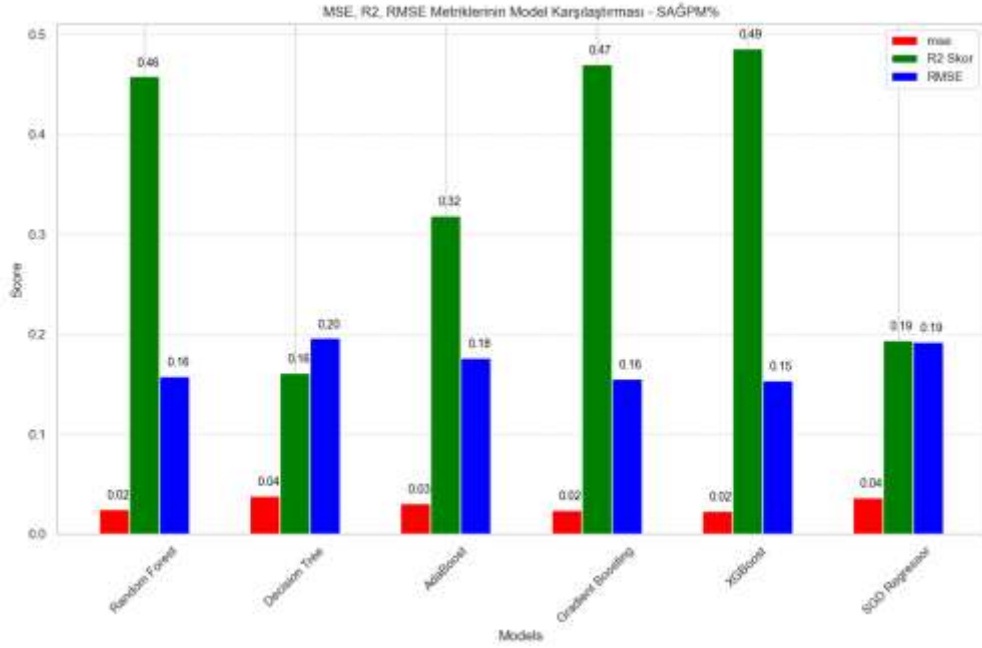
Şekil 5.2. SOLANTR% değeri tahmininde model performansları.

Sonuçlar, SAĞANTR% tahmin sonuçlarına göre gözlemlenen bir kötüleşmeyi ortaya koymaktadır. Tablodaki regresyon modellerinin R2 değerleri -0.35 ile 0.23 arasında değişim göstermektedir. Bu değerler, modellerin tahmin performansının veri setinin varyansının bir bölümünü açıkladığını gösterirken, aynı zamanda iyileştirme fırsatlarının da olduğunu işaret eder. Özellikle, karar ağacı modeli diğer modellere kıyasla daha düşük bir performans sergilemiştir. Bu sonuçlar, veri setindeki özellikler ve karmaşıklıkların daha etkili bir şekilde ele alınması gerektiğini vurgular. Bu bağlamda, modelin hiperparametre ayarları, veri ön işleme adımları ve daha iyi özellik seçimi gibi faktörler gözden geçirilmelidir. Elde edilen sonuçlar, modelin genel performansının artırılması ve daha kesin tahminler elde edilmesi için potansiyel geliştirme alanlarını belirlemek için dikkate alınmalıdır.

Çizelge 5.3'te ve Şekil 5.3'te SAĞPM% tahmini için algoritmaların performansı gösterilmiştir. En iyi sonucu 0,02 MSE, 0,49 R2 ve 0,15 RMSE değeri ile XGBoost algoritması vermiştir.

Çizelge 5.3. SAĞPM% değeri tahmininde regresyon algoritmaları sonuçları.

Değişkenler	Hedef YDT değeri	Metrik	Makine Öğrenimi Regresyon Algoritmaları					
			Decison Tree	Random Forest	AdaBoost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş	SAĞPM%	MSE	0,04	0,02	0,03	0,02	<b>0,02</b>	0,04
		R2	0,16	0,46	0,32	0,47	<b>0,49</b>	0,19
		RMSE	0,20	0,16	0,18	0,16	<b>0,15</b>	0,19



Şekil 5.3. SAĞPM% değeri tahmininde model performansları.

Elde edilen sonuçlar ve analizler doğrultusunda, SAĞPM% değerini tahmin etmek üzere altı farklı regresyon modeli değerlendirilmiştir. Random Forest, Gradient Boosting ve XGBoost algoritmaları benzer şekilde sonuçlar vermiştir. Bu modeller arasında XGBoost'un öne çıktığı görülmektedir. XGBoost'un üstün performans sergilemesinin arkasında birkaç etken yer almaktadır.

Özellikle, özellik seçimi ve önemi belirleme yeteneği sayesinde, daha etkili özelliklerin öne çıkarılması ve daha kesin tahminler elde edilmesi sağlanabilir. Aynı

zamanda ensemble bir yöntem olan XGBoost, birden fazla zayıf tahminiciyi birleştirerek modelin karmaşıklığını artırır ve veri setindeki desenleri daha iyi yakalar. Aşırı öğrenme kontrolü, dengesiz veri setlerini yönetme yeteneği, regülerleme teknikleri ve optimize edilmiş parametreler de XGBoost'un performansını artırıcı faktörlerdir. Bu faktörlerin bir araya gelmesi, XGBoost'un diğer algoritmalar karşısında üstün performans sergilemesine olanak sağlar. Elde edilen sonuçlar, özellikle XGBoost'un SAĞPM% tahmini konusunda daha güvenilir ve kesin tahminler sunduğunu göstermektedir. Bu durum, veri setinin özelliklerine ve hedeflerine bağlı olarak en iyi algoritma seçiminin farklılık gösterebileceğini ortaya koyar.

Çizelge 5.4'te ve Şekil 5.4'te SOLPM% tahmini için algoritmaların performansı gösterilmiştir. En iyi sonucu 0,04 MSE, 0,42 R2 ve 0,14 RMSE değeri ile Decision Tree algoritması vermiştir.

Çizelge 5.4. SOLPM% değeri tahmininde regresyon algoritmaları sonuçları.

Değişkenler	Hedef YDT değeri	Metrik	Makine Öğrenimi Regresyon Algoritmaları					
			Decision Tree	Random Forest	AdaBoost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş	SOL PM%	MSE	<b>0,04</b>	0,03	0,03	0,03	0,03	0,03
		R2	<b>0,42</b>	0,25	0,23	0,26	0,22	0,15
		RMSE	<b>0,14</b>	0,16	0,17	0,16	0,17	0,17

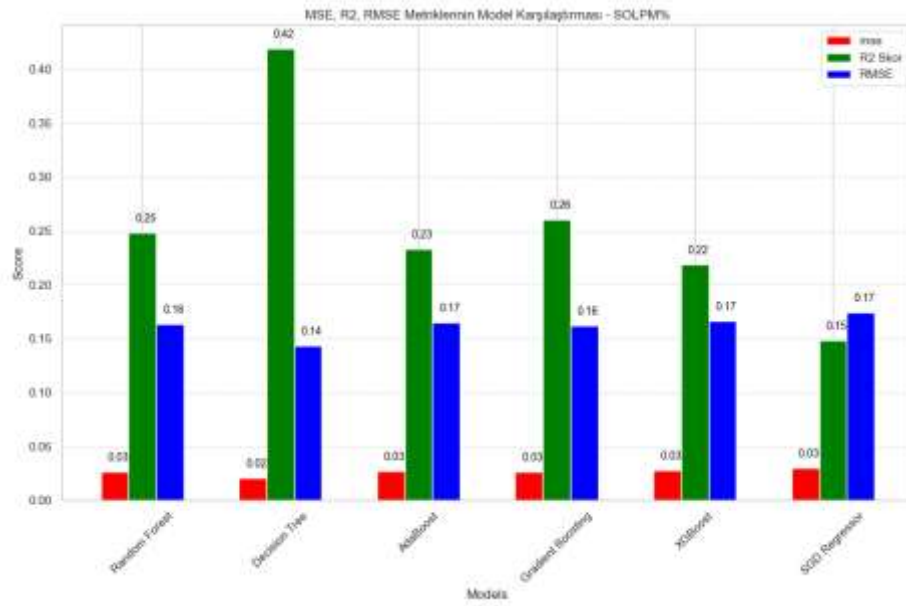
Verilen sonuçlara göre, SOLPM% hedef değişkenini tahmin etme konusunda Decision Tree algoritmasının diğer algoritmalarından daha başarılı olduğu gözlenmektedir. Bu başarı, özellikler arasındaki belirgin ilişkinin yakalanması, özellik seçimi ve ön işleme stratejileri ile ilişkilendirilebilir. Ayrıca, veri dağılımı, modelin aşırı öğrenme eğilimi ve hiperparametre ayarlarının da etkisi olabilir. Bu sonuçlar, özelliklerin ve veri setinin yapısına bağlı olarak Karar Ağacı bazı durumlarda daha iyi genelleme yapabileceğini işaret etmektedir.

Birincil olarak, Karar Ağacı algoritması, özellikler arasındaki doğrusal olmayan ilişkileri yakalama yeteneği ile bilinir. SOLPM% hedef değişkeni için özellikler

arasında karışık veya karmaşık bir ilişki varsa, Karar Ağacı bu ilişkiyi daha etkili bir şekilde modelleyebilir.

İkincil olarak, SOLPM% hedef değişkeninin özelliklerle etkileşimi veya ilişkisi, diğer hedef değişkenlere göre daha belirgin olabilir. Bu da Karar Ağacı algoritmasının bu özel duruma daha iyi adapte olmasını sağlayabilir.

Ayrıca, veri setinin yapısı, SOLPM% hedef değişkeninin dağılımı ve veri noktalarının dağılımı gibi faktörler de sonuçları etkileyebilir. Her bir algoritmanın veri setinin farklı özelliklerini farklı şekillerde ele alması, sonuçların bu şekilde farklılaşmasına neden olmaktadır.



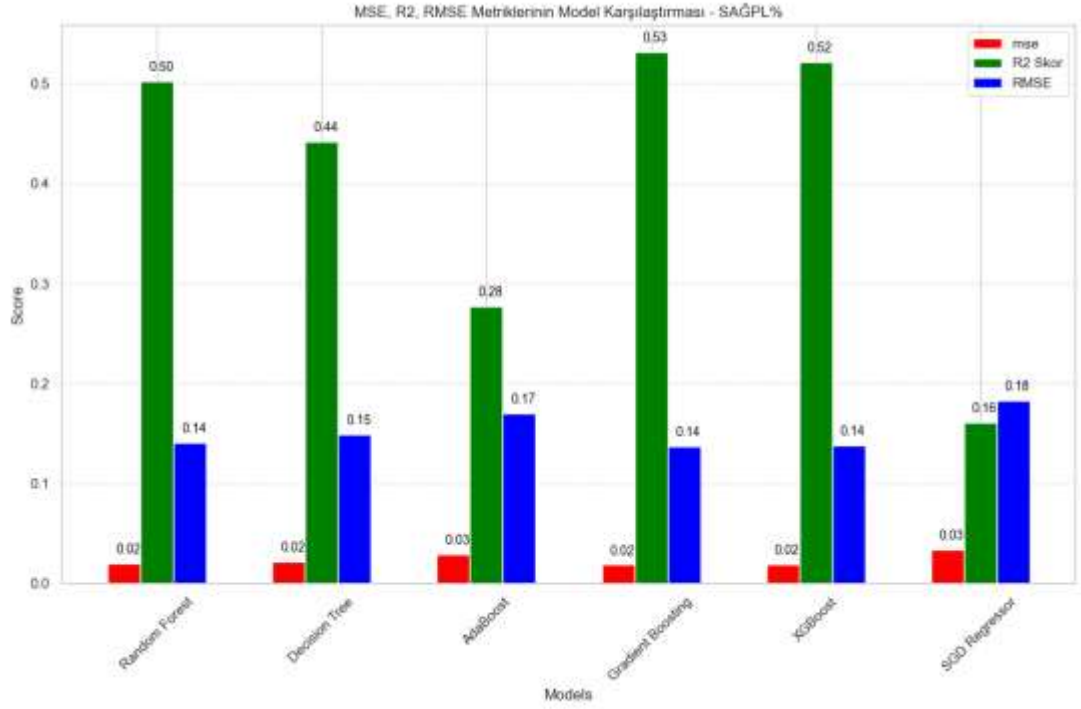
Şekil 5.4. SOLPM% değeri tahmininde model performansları.

Çizelge 5.5'te ve Şekil 5.5'te SAĞPL% tahmini için algoritmaların performansı gösterilmiştir. En iyi sonucu 0,02 MSE, 0,53 R2 ve 0,14 RMSE değeri ile Gradient Boosting algoritması vermiştir.

Çizelge 5.5. SAĞPL% değeri tahmininde regresyon algoritmaları sonuçları.

			Makine Öğrenimi Regresyon Algoritmaları					
Değişkenler	Hedef YDT değeri	Metrik	Decision Tree	Random Forest	AdaBoost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş	SAĞ PL%	MSE	0,02	0,02	0,03	<b>0,02</b>	0,02	0,03
		R2	0,44	0,50	0,28	<b>0,53</b>	0,52	0,16
		RMSE	0,15	0,14	0,17	<b>0,14</b>	0,14	0,18

Sonuçlar, ilk bakışta, Random Forest, Gradient Boosting ve XGBoost gibi ensemble algoritmalarının, aynı zamanda Decision Tree algoritmasının da iyi sonuçlar verdiğini göstermektedir. Bu durum, hem ensemble algoritmaların hem de tekil Karar Ağacı algoritmasının veri setindeki desenleri ve özellikleri etkili bir şekilde modelleyebildiğini işaret eder. Her bir algoritmanın, farklı özellikleri ve yapıları yakalama yeteneği sayesinde, hedef değişkenin tahmininde başarılı sonuçlar elde edebilmiş olduğu görülmektedir. Gradient Boosting algoritmasının yüksek R2 değeri, bu algoritmanın bağımlı değişkeni ne kadar iyi açıkladığını gösterir. Bu da, modelin tahminlerinin gerçek değerlere ne kadar yakın olduğunu yansıtabilir. Bununla birlikte, düşük MSE ve RMSE değerleri de, tahminlerin gerçek değerlere yakın olduğunu ve modelin düşük hata oranıyla çalıştığını gösterir.



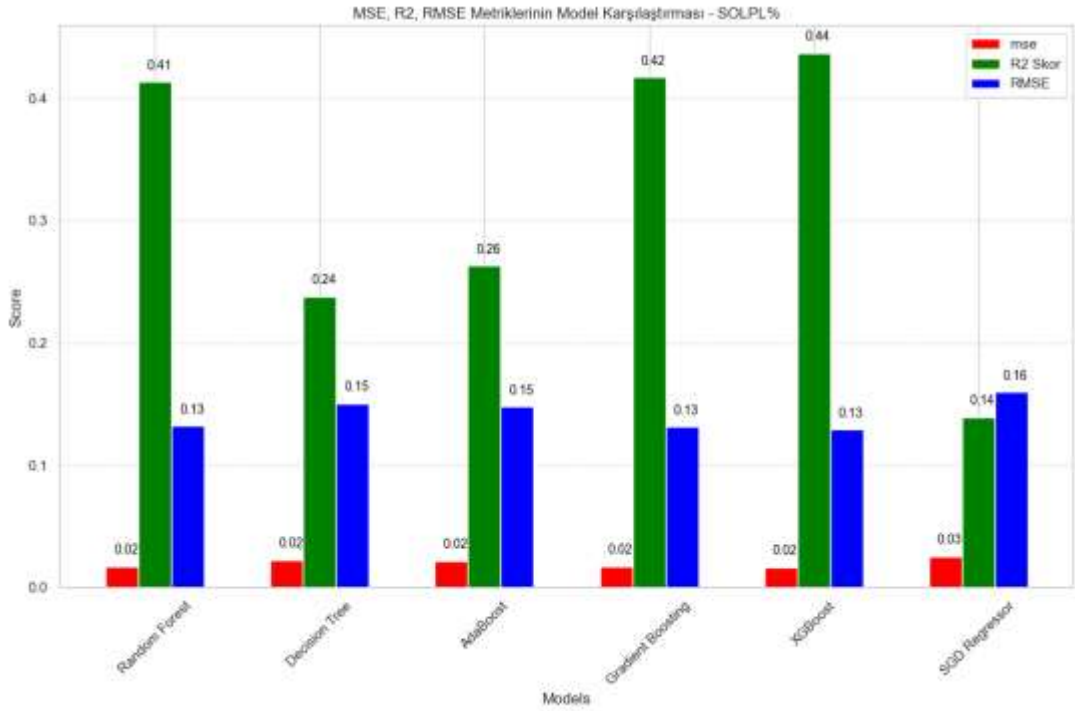
Şekil 5.5. SAĞPL% değeri tahmininde model performansları.

Bu başarı, ensemble bir yöntem olan Gradient Boosting'in birden fazla zayıf tahminciyi birleştirerek güçlü bir model oluşturmasına bağlanabilir.

Çizelge 5.6'da ve Şekil 5.6'da SOLPM% tahmini için algoritmaların performansı gösterilmiştir. En iyi sonucu 0,02 MSE, 0,44 R2 ve 0,13 RMSE değeri ile XGBoost algoritması vermiştir.

Çizelge 5.6. SOLPL% değeri tahmininde regresyon algoritmaları sonuçları.

		Makine Öğrenimi Regresyon Algoritmaları						
Değişkenler	Hedef YDT değeri	Metrik	Decision Tree	Random Forest	AdaBoost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş	SOL PL%	MSE	0,02	0,02	0,02	0,02	<b>0,02</b>	0,03
		R2	0,24	0,41	0,26	0,42	<b>0,44</b>	0,14
		RMSE	0,15	0,13	0,15	0,13	<b>0,13</b>	0,16



Şekil 5.6. SOLPL% d değeri tahmininde model performansları.

Tablodaki SOLPL% değerinin regresyon algoritmalarındaki metrik sonuçları incelendiğinde, farklı algoritmaların SOLPL% hedef değişkenini tahmin etme konusundaki performansı görülmektedir.

MSE değerleri incelendiğinde, Random Forest, Karar Ağacı, AdaBoost, Gradient Boosting ve XGBoost algoritmalarının tümünün düşük hata değerleri elde ettiği gözlemlenmektedir. Bu, algoritmaların tahminlerinin gerçek değerlere ne kadar yakın olduğunu ve hata miktarının düşük olduğunu gösterir. SGD algoritmasının ise biraz daha yüksek bir MSE değeri olduğu fark edilmektedir.

R2 değerleri değerlendirildiğinde, Gradient Boosting ve XGBoost algoritmalarının yüksek R2 değerleri ile öne çıktığı görülmektedir. Bu, bu iki algoritmanın bağımlı değişkeni daha iyi açıkladığını ve tahminlerin gerçek değerlere daha yakın olduğunu gösterir. Diğer algoritmaların da pozitif R2 değerleri ile genel olarak iyi bir tahmin performansı sergilediği görülmektedir.

RMSE deęerleri incelendięinde, tm algoritmaların benzer dzeyde dřk RMSE deęerleri elde ettięi grlmektedir. Bu da algoritmaların tahminlerinin ortalama olarak geręek deęerlere yakın olduęunu ve dřk hata oranına sahip olduęunu gsterir.

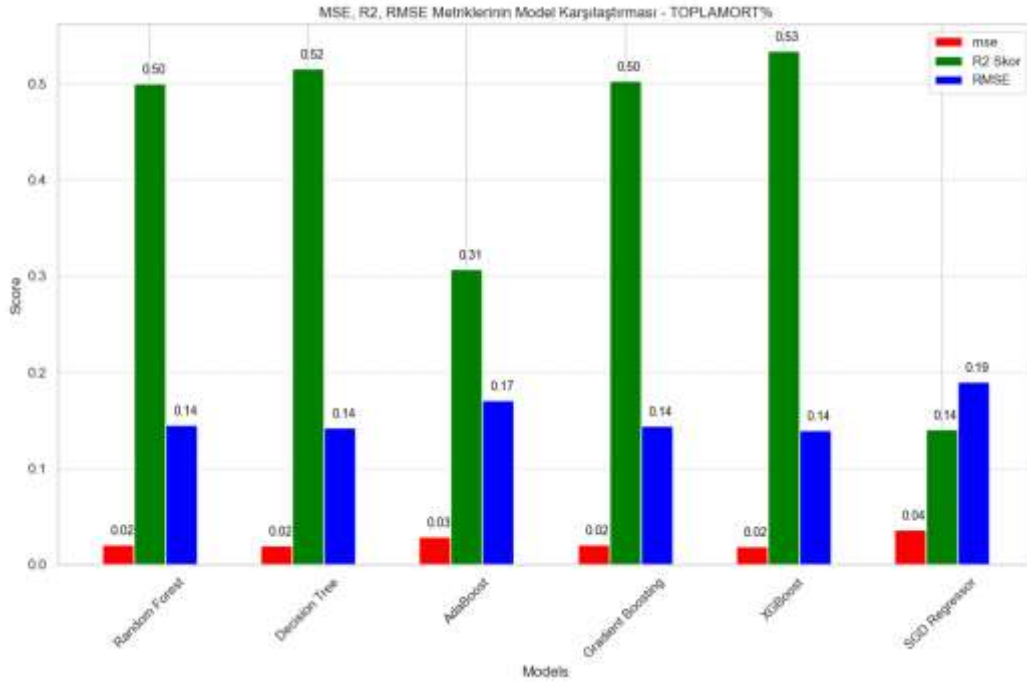
Sonuç olarak, SOLPL% deęerinin regresyon algoritmalarındaki metrik sonuları incelendięinde, Gradient Boosting ve XGBoost algoritmalarının SOLPL% hedef deęiřkenini tahmin etme konusunda daha stn performans sergiledięi gzlemlenmektedir. Dięer algoritmaların da genel olarak iyi bir tahmin performansı gsterdięi ve SOLPL% hedef deęiřkeninin regresyon modelleri ile bařarılı bir řekilde tahmin edilebildięi grlmektedir.

izelge 5.7’de ve řekil 5.7’de TOPLAMORT% tahmini iin algoritmaların performansı gsterilmiřtir. TOPLAMORT% deęerinin tahmininin yapılmasındaki asıl ama denge verilerini yorumlayan ve sporcuları ynlendiren uzmanlara tarafından kiřininin dinamik denge performansı hakkında bilgi vermesidir. Girdi verilerinden ıktı olarak elde ettięimiz bu denge verisi daha hızlı ve yorumlanabilir dinamik denge performansı saęlar. Random Forest, Gradient Boosting ve XGBoost ensemble algoritmalarında bařarılı sonular vermiřtir. En iyi sonucu 0,02 MSE, 0,53 R2 ve 0,14 RMSE deęeri ile XGBoost algoritması vermiřtir.

izelge 5.7. Cinsiyet, Yař, BMI, Antrenman Yılı, Branř deęiřkenleri kullanarak TOPLAMORT% deęerinin regresyon algoritmaları sonuları.

Deęiřkenler	Hedef YDT deęeri	Metrik	Makine ęrenimi Regresyon Algoritmaları					
			Decision Tree	Random Forest	AdaBoost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yař, BMI, Antrenman Yılı, Branř	TOPLAM ORT%	MSE	0,02	0,02	0,03	0,02	<b>0,02</b>	0,04
		R2	0,52	0,50	0,31	0,50	<b>0,53</b>	0,14
		RMSE	0,14	0,14	0,17	0,14	<b>0,14</b>	0,19





Şekil 5.7. TOPLAMORT% değeri tahmininde model performansları.

Çizelge 5.8’de değişkenlerin (Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş) yanı sıra SAĞANT% özelliğini modele eklediğinizde elde edilen sonuçların gözle görülür şekilde yükseldiği görülmektedir. Bu durum, SAĞANT% değişkeninin tahmin performansına büyük katkı sağladığını ve diğer değişkenlerle güçlü bir ilişki içerdiğini göstermektedir. Yani, SAĞANT% değişkeni tahmin modeline eklenerek tahmin gücü artmıştır. Bu sonuç, bu değişkenin hedef değişkeni tahmininde kritik bir rol oynadığını ve modelin daha iyi performans göstermesini sağladığını göstermektedir.

Öncelikle, SAĞANT% hedef değişkeni incelendiğinde, tüm algoritmaların düşük hata değerleri (MSE) elde ettiği gözlenir. Bu, SAĞANT% değerinin tahminlenmesinde algoritmaların başarılı olduğunu gösterir. R2 değerleri incelendiğinde, Gradient Boosting ve XGBoost algoritmalarının yüksek R2 değerleri ile dikkat çektiği görülür. Bu da bu iki algoritmanın hedef değişkeni daha iyi açıklama yeteneğine sahip olduğunu gösterir. RMSE değerleri ise tahminlerin gerçek değerlere yakınlığını ölçerken, tüm algoritmaların benzer düzeyde düşük RMSE değerleri elde ettiği gözlemlenir. SOLANT%, SAĞPM%, SOLPM%, SAĞPL% ve SOLPL% hedef değişkenleri için de benzer şekilde analiz yapılmıştır. Bu hedef değişkenlerinin tahminlerinde de çoğu algoritmanın düşük hata değerleri elde ettiği görülmüştür.

Ancak, hedef deęişkenine göre algoritmaların R2 ve RMSE deęerleri farklılık göstermiştir. Örneęin, SAĞPM% hedef deęişkeni için Gradient Boosting ve XGBoost algoritmalarının daha yüksek R2 deęerleri ile öne çıktığı görülürken, SOLPM% hedef deęişkeni için XGBoost ve AdaBoost algoritmalarının daha iyi tahmin performansı sergilediğı dikkat çekmektedir.

Çizelge 5.8. Cinsiyet, Yaş, BMI, Antrenman Yılı ve Branş deęerlerine ile YDT verisi (SAĞANTR%) verilerek geri kalan YDT deęerlerinin tahmini.

Deęişkenler	Hedef YDT deęeri	Metrik	Makine Öğrenimi Regresyon Algoritmaları					
			Decision Tree	Random Forest	Ada Boost	Gradient Boosting	XG Boost	SGD
Cinsiyet, Yaş, BMI, Antrenman Yılı, Branş, SAĞANT%	SOLANT%	MSE	0,02	<b>0,01</b>	0,01	<b>0,01</b>	0,01	0,03
		R2	0,61	<b>0,77</b>	0,72	<b>0,77</b>	0,67	0,27
		RMSE	0,12	<b>0,09</b>	0,1	<b>0,09</b>	0,11	0,17
	SAĞPM%	MSE	0,03	0,02	0,03	<b>0,02</b>	0,03	0,03
		R2	0,29	0,53	0,43	<b>0,57</b>	0,4	0,29
		RMSE	0,18	0,15	0,16	<b>0,14</b>	0,17	0,18
	SOLPM%	MSE	0,03	0,02	0,02	0,02	<b>0,02</b>	0,03
		R2	0,09	0,34	0,36	0,35	<b>0,39</b>	0,22
		RMSE	0,18	0,15	0,15	0,15	<b>0,15</b>	0,17
	SAĞPL%	MSE	0,02	0,02	0,02	<b>0,02</b>	0,02	0,03
		R2	0,39	0,6	0,5	<b>0,61</b>	0,53	0,32
		RMSE	0,16	0,13	0,14	<b>0,13</b>	0,14	0,16
	SOLPL%	MSE	0,02	0,02	0,02	0,02	<b>0,02</b>	0,02
		R2	0,18	0,45	0,4	0,5	<b>0,46</b>	0,29
		RMSE	0,16	0,13	0,13	0,12	<b>0,13</b>	0,14
	TOPLAM ORT%	MSE	0,02	0,02	0,02	<b>0,01</b>	0,02	0,03
		R2	0,45	0,6	0,57	<b>0,66</b>	0,59	0,32
		RMSE	0,15	0,13	0,13	<b>0,12</b>	0,13	0,17

Sonuç olarak, çizelge üzerinden elde edilen sonuçlar incelendiğinde, regresyon algoritmalarının farklı hedef deęişkenlerini tahmin etme yeteneklerinin deęişkenlik gösterdiği görülmüştür. Bu farklılık, algoritmaların veri setine ve hedef deęişkenine göre adaptasyon yeteneğine baęlı olabilir. Ayrıca, algoritmaların metrik deęerleri üzerinden performanslarının analiz edilmesi, hangi algoritmanın hangi hedef deęişkenlerinde daha iyi tahminler yaptığını belirlemek açısından önemlidir.

## BÖLÜM 6

### SONUÇLAR VE ÖNERİLER

Bu çalışma, spor bilimlerinde kullanılan YDT verilerinin makine öğrenme yöntemleriyle analizini gerçekleştirmeyi amaçlamıştır. Regresyon problemi olarak ele alınan bu çalışmada, farklı makine öğrenimi regresyon algoritmalarının sonuçları değerlendirilmiş ve her bir denge verisinin sonuçlara etkisi gözlemlenmiştir. Veri temizleme, normalizasyon gibi ön işlemlerden geçen veri seti, algoritmaların anlayabileceği bir formata dönüştürülmüş ve daha anlaşılır hale getirilmiştir. Çalışma kapsamında, her bir denge testi verisi için farklı modeller ve algoritmalar kullanılarak detaylı bir analiz gerçekleştirilmiştir. Bu yaklaşım sayesinde, hangi denge verisinin hangi algoritma veya modelde daha iyi sonuç verdiği ve veri setine etkisinin ne olduğu gözlemlenmiştir. Bu çalışma, spor performansı ve sağlık analizinde YDT verilerinin kullanımını daha sağlam temellere oturtma konusunda önemli bir adım olarak değerlendirilebilir.

Makine öğrenimi modellerinin performans değerlendirmesine göre, ensemble modelleri (Random Forest, Gradient Boosting, XGBoost YDT sonuçlarını tahmin etme konusunda daha iyi sonuçlar elde etmiştir. Bu modeller, düşük hata oranları (MSE), yüksek varyans açıklama oranları (R<sup>2</sup>) ve düşük hata dağılımları (RMSE) ile öne çıkmıştır. Özellikle SAĞANT gibi referans bir denge değeri tahmin modellerine eklenerek tahmin performansının artırıldığı gözlemlenmiştir.

Bu tez çalışması, spor bilimlerinde YDT verilerinin makine öğrenme yöntemleri ile analizini gerçekleştirerek olumlu sonuçlar elde etmiştir. Ancak, gelecekteki çalışmalarda veri setinin daha geniş ve çeşitli hale getirilmesi, tahmin modellerinin daha yüksek başarı elde etme potansiyelini artırabilir.

Farklı spor dallarından daha fazla katılımcı içeren veri setleri, özellikle branş sayılarının çeşitlendirilmesi, veri setindeki dengesizliği azaltabilir ve modellerin genelleme yeteneklerini daha iyi değerlendirmemize olanak tanıyabilir. Bu tür bir genişleme, tahmin performansının daha yüksek ve güvenilir seviyelere çıkabileceği umudunu taşımaktadır. Bu nedenle, tez sonuçlarının olumlu görünümüne rağmen, veri setinin artırılmasının daha yüksek sonuçlar elde etme potansiyeli taşıdığı unutulmamalıdır.

Son olarak, bu çalışmanın sonuçları, YDT sonuçlarının sporcuların performansı ve sakatlık riski tahmini gibi alanlarda nasıl kullanılabilceği konusunda daha fazla araştırmayı teşvik etmektedir. Özellikle sporcuların performansını optimize etmek ve sakatlık riskini en aza indirmek için y-denge testi sonuçlarının klinik ve antrenman uygulamalarında nasıl kullanılabilceği daha ayrıntılı bir şekilde incelenebilir. Bu öneriler, gelecekteki çalışmaların daha kapsamlı ve uygulama odaklı sonuçlara ulaşmasını sağlayabilir.

## KAYNAKLAR

1. İncetaş M. O., Uçar M., Bayraktar I. ve Çilli M., “using machine learning algorithms for jumping distance prediction of male long jumpers”, *Journal of Intelligent Systems: Theory and Applications*, 5(2): 145-152 (2022).
2. Akay, M. F., Ates, B., and Ozciloglu, M. M., “Prediction of y-balance test performance using generalized regression neural network”, *In Proceedings of the 1st International Conference on Life and Engineering Sciences*, Kyrenia, Cyprus, 21 (2018).
3. Gil-Martín, M., Johnston, W., San-Segundo, R. and Caulfield, B., “Scoring performance on the y-balance test using a deep learning approach”, *Sensors*, 21 (21): 7110 (2021).
4. Johnston, W., O'Reilly, M., Dolan, K., Reid, N., Coughlan, G. and Caulfield, B., “objective classification of dynamic balance using a single wearable sensor”, *4th International Congress on Sport Sciences Research and Technology Support*, 15-24 (2016).
5. S. Uslu, "Farklı ağırlıklar ile yapılan squat sıçramanın makine öğrenme yöntemleri ile değerlendirilmesi", *Akdeniz Spor Bilimleri Dergisi*, 5 (1): 1-12 (2022).
6. Ruiz-Pérez, I., “A field-based approach to determine soft tissue injury risk in elite futsal using novel machine learning techniques”, *Frontiers in Psychology*, 12 (2021).
7. Oytun, M., Tinazci, C., Sekeroglu, B., Acikada, C. and Yavuz, H. U., "Performance prediction and evaluation in female handball players using machine learning models", *IEEE Access*, 8: 116321-116335, (2020).
8. Jauhiainen, S., Kauppi, J., Leppänen, M., Pasanen, K., Parkkari, J., Vasankari, T., Kannus, P. and Äyrämö, S., “New Machine Learning Approach for Detection of Injury Risk Factors in Young Team Sport Athletes”, *International Journal of Sports Medicine*, 42 (2020).
9. Huang, Y., “The impact of sport-specific physical fitness change patterns on lower limb non-contact injury risk in youth female basketball players: a pilot study based on field testing and machine learning.” *Frontiers in Physiology*, 14: 1182755 (2023).

10. Şimşek, M. ve Kesilmiş, İ., “Predicting athletic performance from physiological parameters using machine learning: example of bocce ball”. *Journal of Sports Analytics*, 8: 1-9 (2022).
11. Robles Palazón, F. J., “Predicting injury risk using machine learning in male youth soccer players”, *Chaos, Solitons & Fractals*, 167 (2023).
12. Manshadi; M., "Postural balance for selection of martial artists using machine learning techniques", *Journal of Exercise and Health Science*, 2 (1): 1-12 (2022).
13. Hrysomallis, C., “Balance Ability and Athletic Performance”. *Sports Med*, 41 (3): 221–232 (2011).
14. Suveren Erdoğan, C., Er, F., İpekoğlu, G., Çolakoğlu, T., Zorba, E., ve Colakoglu, F., “Farklı denge egzersizlerinin voleybolcularda statik ve dinamik denge performansı üzerine etkileri”, *Spor ve Performans Araştırmaları Dergisi*, 8: 11-18 (2017).
15. Plisky, P. J., Gorman, P. P., Butler, R. J., Kiesel, K. B., Underwood, F. B., and Elkins, B., “The reliability of an instrumented device for measuring components of the star excursion balance test”, *North American Journal of Sports Physical Therapy: NAJSPT*, 4 (2): 92-99 (2009).
16. Plisky, P.J., Rauh, M.J., Kaminski, T.W. and Underwood, F.B., “Star Excursion Balance Test as a predictor of lower extremity injury in high school basketball players”, *J Orthop Sports Phys Ther*, 36 (12): 911-919 (2006).
17. Thorpe, J., and Ebersole, K., “unilateral balance performance in female collegiate soccer athletes”, *Journal of Strength and Conditioning Research / National Strength & Conditioning Association*. 22: 1429-33 (2008).
18. Linek, P., Sikora, D., Wolny, T., and Saulicz, E., “Reliability and number of trials of Y Balance Test in adolescent athletes”, *Musculoskeletal Science and Practice*, 31: 72-75 (2017).
19. Gribble, P.A., Hertel, J. and Plisky, P., “Using the star excursion balance test to assess dynamic postural-control deficits and outcomes in lower extremity injury: a literature and systematic review”, *Journal of Athletic Training*, 47: 339–357 (2012).
20. Grace Gaerlan, M., Alpert, P. T., Cross, C., Louis, M. and Kowalski, S., “Postural balance in young adults: the role of visual, vestibular and somatosensory systems”, *Journal of the American Academy of Nurse Practitioners*, 24 (6): 375-381 (2012).

21. Butler, R.J., Southers, C., Gorman, P.P., Kiesel, K.B. and Plisky, P.J., “Differences in soccer players dynamic balance across levels of competition”, *Journal of Athletic Training*, 47 (6): 616–620 (2012).
22. Stepinski, M., Zwierko, T., Florkiewicz, B. and Debicka J., “The level of chosen motor abilities of 13 years old soccer players”, *Journal of Human Kinetics*, 9: 99-106 (2003).
23. Huxham, F. E., Goldie, P. A. and Patla, A. E., “Theoretical considerations in balance assessment”, *Australian Journal of Physiotherapy*, 47 (2): 89-100 (2001).
24. Şimşek, D. ve Ertan, H., “Postural kontrol ve spor: spor branşlarına yönelik postural sensör-motor stratejiler ve postural salınım”, *Spormetre Beden Eğitimi ve Spor Bilimleri Dergisi*, 9 (3): 81-90 (2011).
25. Dick, R., Agel, J. and Marshall, S.W., “National collegiate athletic association injury surveillance system sommentaries: Introduction and methods”, *Journal of Athletic Training*. 42 (2): 173-182 (2007).
26. Hootman, J.M., Dick, R. and Agel J., “Epidemiology of collegiate injuries for 15 sports: summary and recommendations for injury prevention initiatives”, *Journal of Athletic Training*, 42 (2): 311–319 (2007).
27. Phillips, L.H., “Sports injury incidence”, *British Journal of Sports Medicine*, 34 (2): 133-136 (2000).
28. Ates, B., Cetin, E., Yarim, I., "Kadın sporcularda denge yeteneği ve denge antrenmanları", *Gaziantep Üniversitesi Spor Bilimleri Dergisi*, 2: 66-79 (2017).
29. Lai, W. C., “Lower quarter y-balance test scores and lower extremity injury in ncaa division I. athletes”, *Orthopaedic Journal of Sports Medicine*, 5(8): 2325967117723666 (2017).
30. Plisky, P., “Systematic review and meta-analysis of the y-balance test lower quarter: reliability, discriminant validity, and predictive validity”, *International Journal of Sports Physical Therapy*, 16 (5): 1190-1209 (2021).
31. Greenberg, E. T., “Interrater and test-retest reliability of the y balance test in healthy, early adolescent female athletes”, *International Journal of Sports Physical Therapy* 14 (2): 204-213 (2019).
32. Ates, B., “Düzenli pilates egzersizi yapan kadınlar ile sedanter kadınlarda y denge testi performansının karsilastirilmesi”, *Ege Tip Bilimleri Dergisi*, 2: 2-7 (2019).
33. Riemann, B. and George, D., “Limb, sex, and anthropometric factors influencing normative data for the biodex balance system sd athlete single leg stability test.” *Athletic Training & Sports Health Care*, 5: 224-232 (2013).

34. Opstoel, K., “Anthropometric characteristics, physical fitness and motor coordination of 9 to 11 year old children participating in a wide range of sports” *Plos One*, 10 (5) (2015).
35. Plisky, P.J., Gorman, P.P. and Butler, R.J., “The reliability of an instrumented device for measuring components of the star excursion balance test”, *N Am J Sports Phys Ther.*, 4: 92-99 (2009).
36. Engquist, K.D., Smith, C.A. and Chimera, N.J., “Performance comparison of student-athletes and general college students on the functional movement screen and the y balance test”, *Journal of Strength and Conditioning Research*, 29 (8):2296-2303 (2015).
37. Wilson, B. R., Robertson, K. E., Burnham, J. M., Yonz, M. C., Ireland, M. L. and Noehren, B., “The relationship between hip strength and the y balance test”, *Journal of Sport Rehabilitation*, 27 (5): 445-450 (2018).
38. Gonell, A.C., Romero, J.A.P. and Soler, L.M., “Relationship between the y-balance test scores and soft tissue injury incidence in a soccer team”, *International Journal of Sports Physical Therapy*, 10: 955–966 (2015).
39. Myers, H., Christopherson, Z. and Butler, R. J., “Relationship between the lower quarter Y-balance test scores and isokinetic strength testing in patients status post ACL reconstruction”, *International Journal of Sports Physical Therapy*, 13(2), 152 (2018).
40. Butler, R. J., Bullock, G., Arnold, T., Plisky, P. and Queen, R., “Competition-level differences on the lower quarter y-balance test in baseball players”, *Journal of Athletic Training*, 51(12): 997-1002 (2016).
41. Salas-Gomez, D., Fernandez-Gorgojo, M., Sanchez-Juan, P., Bercero, E. L., Perez-Núñez, M. I., and Barbado, D., “Quantifying balance deficit in people with ankle fracture six months after surgical intervention through the Y-Balance test”, *Gait & Posture*, 95: 249-255 (2020).
42. Hallagin, C., Garrison, J. C., Creed, K., Bothwell, J. M., Goto, S. and Hannon, J., “The relationship between pre-operative and twelve-week post-operative y-balance and quadriceps strength in athletes with an anterior cruciate ligament tear”, *International Journal of Sports Physical Therapy*, 12 (6): 986 (2017).
43. Garrison, J.C., Bothwell, J.M., Wolf, G., Aryal, S. and Thigpen, C.A., “Y balance test anterior reach symmetry at three months is related to single leg functional performance at time of return to sports following anterior cruciate ligament reconstruction”, *International Journal of Sports Physical Therapy*, 10(5): 602–11 (2015).



44. Clagg, S., Paterno, M.V., Hewett, T.E. and Schmitt L.C., “Performance on the modified star excursion balance test at the time of return to sport following anterior cruciate ligament reconstruction”, *The Journal of Orthopaedic & Sports Physical Therapy*, 45(6): 444-452 (2015).
45. Dudek, K., Drużbicki, M., Przysada, G. and Śpiewak, D., “Assessment of standing balance in patients after ankle fractures”, *Acta of Bioengineering and Biomechanics*, 16(4): 59-65 (2014).
46. Earl, J.E. and Hertel, J., “Lower-extremity muscle activation during the star excursion balance tests”, *Journal of Sport Rehabilitation*, 10: 93-104 (2001).
47. Neves, L. F., Souza, C. Q. D., Stoffel, M. and Picasso, C. L. M., “The y balance test—how and why to do it”, *International Physical Medicine & Rehabilitation Journal*, 2(4): 261-262 (2017).
48. Powden, C. J., Dodds, T. K. and Gabriel, E.H., “The reliability of the star excursion balance test and lower quarter y-balance test in healthy adults: a systematic review”, *International Journal of Sports Physical Therapy*, 14(5): 683-694 (2019).
49. Coughlan, G.F., Fullam, K., Delahunt, E., Gissane, C. and Caulfield, B.M., “A comparison between performance on selected directions of the star excursion balance test and the Y balance test”, *Journal of Athletic Training*, 47(4): 366-371 (2012).
50. Linek, P., Sikora, D., Wolny, T. and Saulicz, E., “Reliability and number of trials of y balance test in adolescent athletes”, *Musculoskeletal Science and Practice*, 31: 72-75 (2017).
51. Munro, A.G. and Herrington, L.C., “Between-session reliability of the star excursion balance test”, *Physical Therapy in Sport*, 11(4): 128-132 2010).
52. O'Connor, S., “Can the y balance test identify those at risk of contact or non-contact lower extremity injury in adolescent and collegiate gaelic games?”, *Journal of Science and Medicine in Sport*, 23 (10): 943-948 (2020).
53. Hébert-Losier, K., “Clinical implications of hand position and lower limb length measurement method on y-balance test scores and interpretations”, *Journal of Athletic Training*, 52(10): 910-917 (2017).
54. Muehlbauer, T., Schwiertz, G., Brueckner, D., Kiss, R., Panzer, S., “Limb differences in unipedal balance performance in young male soccer players with different ages”, *Sports*, 7: 20 (2019).
55. Hastie, T., Tibshirani, R. and Friedman, J., “The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition”, *Springer Series in Statistics*, (2019).

56. İnternet: Turhost, “Makine Öğrenmesi (Machine Learning) Nedir?”  
<https://www.turhost.com/blog/makine-ogrenmesi-machine-learning-nedir/#serp>
57. Hastie, T., Tibshirani, R. and Friedman, J., “The elements of statistical learning: data mining, inference, and prediction, second edition”, *Springer Series in Statistics*, (2019).
58. Chapelle, O., Schölkopf, B. and Zien, A., “Semi-supervised learning”, *MIT press*, 2 (2006).
59. Kaelbling, L. P., Littman, M. L., and Moore, A. W., “Reinforcement learning: a survey”, *Journal of Artificial Intelligence Research*, 4: 237-285 (1996).
60. Kantardzic, M., “Data mining: concepts, models, methods, and algorithms 3rd edition”, *Wiley-IEEE Press*, (2019).
61. İnternet: Sckit-Learn, “Decision Tree Reression”,  
[https://scikitlearn.org/stable/auto\\_examples/tree/plot\\_tree\\_regression.html](https://scikitlearn.org/stable/auto_examples/tree/plot_tree_regression.html)
62. Breiman, L., “Random Forests”, *Machine Learning*, 45: 5–32 (2001).
63. İnternet: GeeksforGeeks, “Random Forest Regression”,  
<https://www.geeksforgeeks.org/random-forest-regression-in-python/>
64. Otchere, D. A., Ganat, T. O. A., Ojero, J. O., Tackie-Otoo, B. N. and Taki, M. Y., “Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions”, *Journal of Petroleum Science and Engineering*, 208 (2022).
65. İnternet: GeeksforGeeks, “Gradient Boosting Regression”,  
<https://www.geeksforgeeks.org/ml-gradient-boosting/>
66. Chen, T. and Carlos, G., "Xgboost: a scalable tree boosting system", *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, (2016).
67. Wojtowysch, S. “Stochastic Gradient Descent with Noise of Machine Learning Type Part I: Discrete Time Analysis”, *Journal of Nonlinear Science*, 33: 45 (2023).
68. Jain, M. and Kumar, V., “A survey of techniques for data cleaning. in data mining: foundations and intelligent paradigms”, *Springer*, 3: 1-37 (2010).

69. Hsu, C. W. and Lin, C. J., “A comparison of methods for multi-class support vector machines”, *IEEE Transactions on Neural Networks*, 13 (2): 415-425 (2002).
70. Wang, X. and Yao, X., “Diversity analysis on imbalanced data sets by using ensemble models in data mining and knowledge discovery for big data”, *Springer*, 394-405 (2008).
71. Runkler, T. A., “Data mining and business intelligence: a guide to productivity”, *John Wiley & Son*, (2007).
72. Internet: Lazy Predict, <https://lazypredict.readthedocs.io/en/latest/>

## ÖZGEÇMİŞ

Süheda AKDAĞ ilkokul ve lise öğrenimini Ankara'da tamamladı. Ufuk Arslan Anadolu Lisesi sayısal bölümden mezun olduktan sonra 2016 yılında Karabük Üniversitesi Bilgisayar Mühendisliği (İngilizce) programına başladı. 2021 yılında lisans eğitimini tamamladıktan sonra aynı yıl başladığı yüksek lisans eğitimini Karabük Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı altında sürdürmektedir. Lisans eğitiminde Programlama Dilleri 1-2 derslerinde "Laboratuvar Asistanı" olarak 2 dönem görev almıştır. 2020 yılı itibarıyla Karabük Üniversitesi Teknoloji Geliştirme Bölgesi'nde yer alan SimurgAI Yazılım Tasarım ve Danışmanlık Hizmetleri Limited Şirketi'nde Yapay Zeka Geliştirici olarak görev yaptı. 2021 yılında Ankara'da bulunan Redevio AR-GE şirketinde "Yapay Zeka Mühendisi" olarak çalıştıktan sonra yine 2021 yılında Ankara ODTÜ Teknokent'te bulunan SmartAlpha A.Ş şirketinde "Makine Öğrenimi Mühendisi" olarak görev almıştır.