

T.C.
TRAKYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE YÜRÜYÜŞ VERİLERİ
KULLANILARAK PARKİNSON HASTALIĞININ ÖN TESPİTİ**

SİNEM AYNA ALTUNTAŞ

YÜKSEK LİSANS TEZİ

HESAPLAMALI BİLİMLER ANABİLİM DALI

Tez Danışmanı: DOÇ. DR. OĞUZHAN ERDEM

EDİRNE-2022

SİNEM AYNA ALTUNTAŞ'ın hazırladığı “**MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE YÜRÜYÜŞ VERİLERİ KULLANILARAK PARKİNSON HASTALIĞININ ÖN TESPİTİ**” başlıklı bu tez, tarafımızca okunmuş, kapsam ve niteliği açısından **HESAPLAMALI BİLİMLER** Anabilim Dalında bir **Yüksek lisans tezi** olarak kabul edilmiştir.

Jüri Üyeleri (Ünvan, Ad, Soyad):

İmza

Doç. Dr. Oğuzhan ERDEM

.....

Dr. Öğr. Üyesi Özlem UÇAR

.....

Dr. Öğr. Üyesi H. Feza CARLAK

.....

Tez Savunma Tarihi: 17/01/2022

Bu tezin Yüksek Lisans tezi olarak gerekli şartları sağladığımı onaylarım.

İmza

Doç. Dr. Oğuzhan ERDEM

Tez Danışmanı

.....

Dr. Öğr. Üyesi Sezer ULUKAYA

İkinci Tez Danışmanı

.....

Trakya Üniversitesi Fen Bilimleri Enstitüsü onayı

.....
Prof. Dr. Hüseyin Rıza Ferhat KARABULUT
Fen Bilimleri Enstitüsü Müdürü

T.Ü.FEN BİLİMLERİ ENSTİTÜSÜ
HESAPLAMALI BİLİMLER ANABİLİM DALI
YÜKSEK LİSANS PROGRAMI
DOĞRULUK BEYANI

Trakya Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada, tüm verilerin bilimsel ve akademik kurallar çerçevesinde elde edildiğini, kullanılan verilerde tahrifat yapılmadığını, tezin akademik ve etik kurallara uygun olarak yazıldığını, kullanılan tüm literatür bilgilerinin bilimsel normlara uygun bir şekilde kaynak gösterilerek ilgili tezde yer aldığını ve bu tezin tamamı ya da herhangi bir bölümünün daha önceden Trakya Üniversitesi ya da farklı bir üniversitede tez çalışması olarak sunulmadığını beyan ederim.

17 / 01 / 2022 .

Sinem AYNA ALTUNTAŞ

Yüksek Lisans Tezi

MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE YÜRÜYÜŞ VERİLERİ
KULLANILARAK PARKİNSON HASTALIĞININ ÖN TESPİTİ

T.Ü. Fen Bilimleri Enstitüsü

Hesaplamalı Bilimler Anabilim Dalı

ÖZET

Parkinson hastalığı bilinen bir tedavisi olmayan, nörolojik bir hastalıktır. Hastalığın erken tespiti ilerleyişinin yavaşlatılması ve etkilerinin en aza indirilebilmesi için önemlidir. Bu sebeple hekimlere yardımcı sistemler geliştirilmesi yönünde çalışmalar yapılmaktadır.

Bu tez çalışmasında yalnızca yürüyüş verileri kullanılarak Parkinson hastası-sağlıklı ayrımı yapabilecek bir sınıflandırıcı tasarlanması amaçlanmıştır. Bu amaç doğrultusunda makine öğrenmesi yöntemleri kullanılarak farklı normalizasyon, çapraz doğrulama ve sınıflandırma algoritmaları üzerine çalışılmıştır. Yapılan bu farklı çalışmaların sonuçları karşılaştırılarak en iyi performansı gösteren yöntem belirlenmiştir.

Kullanılan yürüyüş verileri ve geliştirilen sistem vasıtasıyla Parkinson hastalarının yalnızca yürüyüş verileri kullanılarak teşhisin yapılmasına yardımcı olunacaktır.

Yıl : 2022

Sayfa Sayısı : 120

Anahtar Kelimeler : Parkinson Hastalığı, Yürüyüş, Makine Öğrenmesi, Sınıflandırma, Yürüyüş Analizi, Öznitelik Çıkarımı, Öznitelik Seçimi

Master's Degree Thesis

PRELIMINARY DETECTION OF PARKINSON'S DISEASE USING MACHINE
LEARNING TECHNIQUES AND GAIT DATA

Trakya University Institute of Natural and Applied Sciences Computational Science
Department

ABSTRACT

Parkinson's disease is a neurological disease with no known cure. Early detection of the disease is important in slowing its progression and minimizing its effects. For this reason, studies are carried out to develop assistive systems for physicians.

In this thesis, it is aimed to design a classifier that can distinguish between Parkinson's patients and healthy people using only gait data. For this purpose, different normalization, cross validation and classification algorithms were studied by using machine learning methods. The results of these different studies were compared and the method with the best performance was determined.

The gait data used and the developed system will help diagnose Parkinson's patients using only gait data.

Year :2022

Number of Pages : 120

Keywords : Parkinson's Disease, Gait, Machine Learning, Classification, Gait Analysis, Feature Extraction, Feature Selection

ÖNSÖZ

Bu çalışma, Trakya Üniversitesi Fen Bilimleri Enstitüsü Hesaplamalı Bilimler Anabilim Dalı'nda yapılan “Makine Öğrenmesi Teknikleri İle Yürüyüş Verileri Kullanılarak Parkinson Hastalığının Ön Tespiti” isimli yüksek lisans tez çalışmasını içermektedir.

Yüksek lisans öğrenimim sırasında destek ve emeklerini esirgemeyen, yoğunluklarına rağmen her zaman bana zaman ayıran, saygıdeğer danışman hocalarım; Sayın Doç. Dr. Oğuzhan ERDEM'e ve Sayın Dr. Öğr. Üyesi Sezer ULUKAYA'ya saygı ve teşekkürlerimi sunarım.

Bu tez çalışmasının temelini oluşturan veri setini paylaşan Malezya MARA Teknoloji Üniversitesi'nden Prof. Nooritawati Md Tahir ve ekibine katkılarından dolayı teşekkür ederim.

Birlikte öğrenmekten büyük keyif aldığım bu tez çalışmamı geliştirmemde yardımcı olan sevgili arkadaşlarım İlke KURT ve Bahar DEMİR'e ve yoğun çalışma temposu içinde akademik çalışmalar için beni her zaman yüreklendiren Trakya Üniversitesi Edirne Teknik Bilimler Meslek Yüksekokulu hocalarıma ve çalışma arkadaşlarıma anlayış ve destekleri için şükranlarımı sunarım.

Hayatım boyunca desteklerini her zaman hissettiğim sevgili annem Sultan AYNA ve babam Hüseyin AYNA'ya ve yoğun çalışmalarım sırasında sabır ve desteğinden dolayı değerli eşim Erhan ALTUNTAŞ'a teşekkür ve sevgilerimi sunarım.

İÇİNDEKİLER

ÖZET.....	iv
ABSTRACT.....	v
ÖNSÖZ.....	vi
İÇİNDEKİLER.....	vii
ŞEKİLLER DİZİNİ.....	x
ÇİZELGELER DİZİNİ.....	xii
DENKLEMLER.....	xiv
SİMGELER VE KISALTMALAR DİZİNİ.....	xvii
BÖLÜM 1.....	1
GİRİŞ.....	1
BÖLÜM 2.....	5
PARKİNSON HASTALIĞI VE YÜRÜYÜŞE ETKİSİ.....	5
2.1. Parkinson Hastalığı Hakkında Genel Bilgiler.....	5
2.1.1. Parkinson Hastalığının Belirtileri.....	7
2.2. Yürüyüş ve Denge.....	11
2.2.1. Yürüyüş Anatomisi.....	13
2.3. Yürüyüş, Denge ve Parkinson Hastalığı.....	17
2.4. Parkinson Hastalığının Tespitine Yönelik Yapılan Genel Çalışma Örnekleri.....	22
2.5. Parkinson Hastalığının Yürüyüş ile Tespitine Yönelik Çalışma Örnekleri.....	27
BÖLÜM 3.....	36
MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE YÜRÜYÜŞ VERİLERİ KULLANILARAK PARKİNSON HASTALIĞININ TESPİTİ.....	36
3.1. Makine Öğrenmesi Hakkında Genel Bilgiler.....	36
3.2. Veri Kümelerinin Elde Edilmesi.....	38
3.3. Öznitelik Çıkarımı.....	39
3.3.1. Minimum Değer.....	42
3.3.2. Maximum Değer.....	42

3.3.3. Ortalama Değer	42
3.3.4. Medyan (Orta) Değer	43
3.3.5. Standart sapma	43
3.3.6. Değer Aralığı.....	43
3.3.7. Entropi.....	43
3.3.8. Çarpıklık Değeri.....	43
3.3.9. Basıklık Değeri.....	44
3.3.10. Kolmogorov	44
3.3.11. Higuchi Fraktal Boyutu (Higuchi Fractal Dimension).....	44
3.3.12. Katz Fraktal Boyutu (Katz Fractal Dimension)	44
3.3.13. Sample entropi	44
3.3.14. Multiscale Sample Entropi (MSE-Çok Ölçekli Entropi)	45
3.3.15. Harmonik Ortalama Değeri.....	45
3.3.16. RMS (Root Mean Square-Karesel Ortalamanın Karekökü) Değeri	45
3.3.17. Varyans	46
3.3.18. Trim Ortalama Değeri (Kırılmış Ortalama Değer)	46
3.3.19. Min-k Değeri.....	46
3.3.20. Max-k Değeri	46
3.3.21. Varyasyon Katsayısı (Coefficient Varyasyon).....	47
3.3.22. Approximate Entropi.....	47
3.3.23. Fast Approximate Entropi	47
3.3.24. Kovaryans (Covariance)	47
3.3.25. DFA (Detrended Fluctuation Analysis-Eğilimden Arındırılmış Dalgalanma Analizi).....	48
3.4. Normalizasyon Yöntemleri	48
3.4.1. Sıfır-Ortalama Normalizasyon	49
3.4.2. Min-Maks Normalizasyon	49
3.5. Öznitelik Seçme Yöntemleri	49
3.5.1. LASSO (Least Absolute Shrinkage and Selection Operator- En Küçük Mutlak Daralma ve Seçme Operatörü)	50
3.5.2. Minimum Artıklık Maksimum İlişki (mRmR)	51
3.5.3. RELIEF	52
3.6. Sınıflandırma Yöntemleri.....	53
3.6.1. SVM (Support Vector Machines/Destek Vektör Makineleri)	53
3.6.2. KNN (K-Nearest Neighbors/K-En Yakın Komşular).....	54

3.6.3 Naive Bayes	56
3.7. Çapraz Doğrulama Yöntemleri	56
3.7.1. Birini Dışarda Bırak Çapraz Doğrulama (Leave-one-out cross validation)...	57
3.7.2. KFold Çapraz Doğrulama (KFold Cross Validation)	57
3.8. Değerlendirme Ölçütleri.....	58
3.8.1. Doğruluk (Accuracy)	59
3.8.2. Hassasiyet (Sensitivity).....	59
3.8.3. Özgüllük (Specificity).....	59
3.8.4. Kesinlik (Precision).....	60
3.8.5. Recall (Hatırlama-Geri Çağırma).....	60
3.8.6. F-ölçütü	60
3.8.7. Gmean	60
3.8.8. Matthews Korelasyon Katsayısı (Mcc Değeri)	60
BÖLÜM 4	62
DENEYSEL SONUÇLAR	62
4.1. Normalizasyon Yöntemlerinin Etkisi.....	63
4.2. Çapraz Doğrulama Yöntemlerinin Kıyaslanması	67
4.3. Sınıflandırma Yöntemlerinin Kıyaslanması.....	71
4.4. Öznitelik Seçme Algoritmalarının Uygulanması	73
4.4.1. LASSO (Least Absolute Shrinkage and Selection Operator- En Küçük Mutlak Daralma ve Seçme Operatörü) Yöntemi İle Öznitelik Seçimi	77
4.4.2. mRmR (Minimum Artıklık Maksimum İlişki) İle Öznitelik Seçimi	78
4.4.3. RELIEF İle Öznitelik Seçimi	79
4.5. Öznitelik Seçme Algoritmalarının Sınıflandırma Sonuçlarına Katkılarının Karşılaştırılması	81
4.6. Öznitelik Seçme Algoritmalarından Elde Edilen Ortak Öznitelikler ve Sınıflandırıcılara Etkileri.....	85
4.6.1. Normalizasyon yapılmadan seçilen ortak öznitelikler ve sınıflandırıcılardaki etkinlikleri	86
4.6.2. Min-Maks Normalizasyon Yöntemi İle Seçilen Ortak Öznitelikler Ve Sınıflandırıcılardaki Etkinlikleri	87
BÖLÜM 5	90
SONUÇ ve TARTIŞMA.....	90
KAYNAKLAR	95
ÖZGEÇMİŞ	101

ŞEKİLLER DİZİNİ

Şekil 2.1. Parkinson hastası olan ve olmayan bireylerin beyinlerindeki Substantia nigra farklılıkları (Hemensaglık, 2021).....	6
Şekil 2.2. Dopamin'in Parkinson hastalığına etkisi (Gözellik, 2018).....	8
Şekil 2.3. Parkinson hastalığının bazı belirtileri (Gözellik, 2018).....	10
Şekil 2.4. Saunders ve ark. (1953) tarafından tanımlanan 6 temel yürüyüş belirleyici (de Pina Filho, Dutra & Raptopoulos, 2005).	13
Şekil 2.5. Yürüyüş döngüsünde ağırlık merkezinin hareketi ile ilgili iki farklı teoriye (Saunders vd., 1953 ve Cavagna vd., 1966) ait görsel (Kuo, 2007).	14
Şekil 2.6. Gravite (vücut ağırlık merkezi) hattı (Özer, Kaner & Bozdogan, 2014).....	15
Şekil 2.7. Yürüme döngüsü ve fazları (Pehlivan, Tunca, Salur & Ersoy, 2017).	16
Şekil 2.8. Yürüyüşün incelenmesinde kullanılan ve kablosuz veri aktarımı sağlayan bir kayıt cihazı örneği (Hasomed, 2021).	20
Şekil 2.9. Yürüyüş analizi sistemlerinde 3 boyutlu kayıt yapan kamera tabanlı sistemlere örnek (mar-systems, 2021).....	21
Şekil 2.10. Ayak tabanı kuvvet verilerinin elde edilmesinde kullanılan sensörlü tabanlık sistemi çıktıları, kullanılışı ve görseli (Medicaptureurs, 2021).....	22
Şekil 2.11. Ayakkabı tabanlığına yerleştirilen 50 Hz frekanslı, 99 farklı sensör bulunan tabanlık-Pedar marka (Novel, 2021).	28
Şekil 2.12. Ayak tabanı basınç bölgelerinin 10 farklı anatomik bölgeye ayrılması (Kimmeskamp & Hennig, 2001).....	28
Şekil 2.13. Hasta ve sağlıklı bireylerin ayak basınç noktaları ortalamalarının karşılaştırılması (Kimmeskamp & Hennig, 2001).	29

Şekil 2.14. Sağlıklı ve Parkinson hastası bireylerin kuvvet sinyalinin güç spektral yoğunluğu karşılaştırması (Abdulhay vd., 2018).....	31
Şekil 3.1. Bu tez çalışmasının makine öğrenim süreci.....	38
Şekil 3.2. Sırasıyla sağlıklı (a) ve hasta (b) bireylerin kinetik verilerinin ortalama değerleri grafiksel olarak verilmiştir.....	42
Şekil 3.3. SVM sınıflandırıcıda iki sınıfın ayrılmasında oluşturulmuş bir hiper düzlem görseli (Ayhan & Erdoğan, 2014).....	54
Şekil 3.4. KNN sınıflandırıcı modeli temsili şeması (Altunkaynak, Başakın & Kartal, 2020).	55
Şekil 3.5. Kfold çapraz doğrulama yönteminde test işlemi ve doğruluk hesabı örneği (Veribilimiokulu, 2021)	58
Şekil 4.1. Çapraz doğrulamalı Min-Maks normalizasyon ile LASSO seçiminde MSE değeri ve grafiği.	78
Şekil 4.2. RELIEF öznitelik seçim yöntemi ile 152 öznitelik ağırlıklandırılmasına dair grafik	80
Şekil 4.3. SVM sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları.	81
Şekil 4.4. KNN sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Leaveout çapraz doğrulama ile).	83
Şekil 4.5. KNN sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Kfold (K=10) çapraz doğrulama ile).	83
Şekil 4.6. Naive Bayes sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Leaveout çapraz doğrulama ile).	84
Şekil 4.7. Naive Bayes sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Kfold (K=5) çapraz doğrulama ile).	85

ÇİZELGELER DİZİNİ

Çizelge 2.1. Parkinson hastalığının tespitinde farklı verilerler yapılmış çalışma örnekleri özeti.	25
Çizelge 2.2. Bu çalışmanın da temelini oluşturan Prof. Nooritawati Md Tahir ve ekibinin yürüyüş verileri hakkında detaylar.....	33
Çizelge 3.1. Bu çalışmada kullanılan veri kümeleri hakkında bilgiler.	39
Çizelge 3.2. Yürüyüş verileri ve çıkarılan öznitelikler.	40
Çizelge 3.3. n adet veriye sahip bir veri setinin birini dışarıda bırakma çapraz doğrulama yöntemi ile çapraz doğrulanması ve sistem performansı hesabı.	57
Çizelge 3.4. Sınıflandırıcı sonuçlarının değerlendirme ölçütlerinin hesaplanmasında kullanılan karmaşıklık matrisi.....	58
Çizelge 4.1. Öznitelik çıkarımı öncesi ham veri ile elde edilen sınıflandırıcı doğrulukları ve normalizasyon yöntemlerinin etkisi.	63
Çizelge 4.2. SVM sınıflandırıcıdan elde edilen farklı sonuçlar (Ham veri ile).	64
Çizelge 4.3. KNN Sınıflandırıcı parametreleri ve ham veriden elde edilen sonuçlar.	65
Çizelge 4.4. 152 öznitelik ile SVM sınıflandırıcıdan elde edilen sonuçlar.....	67
Çizelge 4.5. KFold ve Leaveout çapraz doğrulama yöntemlerinin ham veri üzerinden KNN sınıflandırıcıda karşılaştırılması.	69

Çizelge 4.6. KNN sınıflandırıcıda Kfold ve Leaveout çapraz doğrulama yöntemlerinin 152 öznitelik kullanılarak karşılaştırılması.	70
Çizelge 4.7. Naive Bayes Sınıflandırma yönteminde 152 öznitelik ile yapılan sınıflandırma sonuçlarında çapraz geçerleme yöntemlerinin karşılaştırılması	71
Çizelge 4.8. Leaveout çapraz doğrulama ile 152 öznitelik kullanılarak elde edilen sınıflandırma en yüksek doğruluk oranları.	71
Çizelge 4.9. Min-Maks normalizasyon ve Leaveout çapraz doğrulama ile 152 öznitelik kullanılarak elde edilen en yüksek sınıflandırma performansları.	72
Çizelge 4.10. 21 öznitelik ve 152 öznitelik ile elde edilen sınıflandırma performanslarının karşılaştırılması.....	74
Çizelge 4.11. 152 öznitelik arasından belirleyici olan ilk 15 özneliğin farklı yöntemlerle belirlenmesi.	75
Çizelge 4.12. Normalizasyon uygulanmadan yapılan seçim işlemlerindeki ortak öznitelikler ve seçen yöntemlerin tablo şeklinde gösterimi.	86
Çizelge 4.13. Normalizasyon yapılmadan seçilen ortak öznitelikler ve farklı sınıflandırıcılardaki performansları.....	87
Çizelge 4.14. Min-Maks Normalizasyon uygulanarak yapılan seçim işlemlerindeki ortak öznitelikler ve seçen yöntemlerin tablo şeklinde gösterimi.	88
Çizelge 4.15. Min-Maks normalizasyon yöntemi ile LASSO, mRmR ve RELIEF yöntemlerinden seçilen ortak öznitelikler ve farklı sınıflandırıcılardaki performansları.	89

DENKLEMLER

Denklem 2.1 Yürüyüş hızının hesaplanması.....	14
Denklem 2.2. GRF Normalizasyon.....	34
Denklem 2.3. İnter grup normalizasyon (grup içi normalizasyon).....	34
Denklem 2.4. İnter grup normalizasyon (gruplar arası normalizasyon).....	34
Denklem 3.1.Ortalama değer hesaplama genel formülü.....	42
Denklem 3.2. Standart sapma hesaplama genel formülü.....	43
Denklem 3.3. Harmonik ortalama hesaplama genel formülü.....	45
Denklem 3.4. RMS değerinin hesaplanmasında kullanılan genel formül.....	46
Denklem 3.5. Varyans değerinin hesaplanmasında kullanılan genel formül.....	46
Denklem 3.6. Varyasyon katsayısının hesaplanmasında kullanılan genel formül.....	47
Denklem 3.7. Kovaryansın hesaplanmasında kullanılan genel formül.....	47
Denklem 3.8. Sıfır ortalama normalizasyon yönteminin genel formülü.....	49

Denklem 3.9. Min-Maks normalizasyon yönteminde bir değeri [0,1] aralığına normalize etmede kullanılan genel formül.....	49
Denklem 3.10. Min-Maks normalizasyon yönteminde bir değeri [0,1] aralığının dışında bir aralığa normalize etmede kullanılan genel formül.....	49
Denklem 3.11. X ve Y öznitelikleri arasındaki benzerlik miktarını ifade eden değerin hesaplanması.....	51
Denklem 3.12. Minimum artıklık (W) değerinin hesaplanmasında kullanılan formül (Gülgezen vd.).....	52
Denklem 3.13. Maksimum ilişki (V) değerinin hesaplanmasında kullanılan formül (Gülgezen vd.).....	52
Denklem 3.14. Ortak bilgi farkı (MID) değerinin hesaplanmasında kullanılan formül ..	52
Denklem 3.15. Ortak bilgi oranı (MIQ) değerinin hesaplanmasında kullanılan formül..	52
Denklem 3.16. RELIEF yönteminin formülü	52
Denklem 3.17. İkili ve devamlı matrisler için uzaklık hesaplama formülleri.....	53
Denklem 3.18. Naive Bayes sınıflandırıcının olasılık hesabı	56
Denklem 3.19. İndirgenmiş Bayes olasılık formülü (Silahdaroğlu, 2016)	56
Denklem 3.20. Karmaşıklık matrisinden doğruluk hesaplanması	59
Denklem 3.21. Karmaşıklık matrisinden hassasiyet hesaplanması.....	59
Denklem 3.22. Karmaşıklık matrisinden özgüllük hesaplanması	59
Denklem 3.23. Karmaşıklık matrisinden kesinlik hesaplanması	60
Denklem 3.24. Karmaşıklık matrisinden recall hesaplanması	60
Denklem 3.25. Karmaşıklık matrisinden F-ölçütü hesaplanması	60

Denklem 3.26. Karmaşıklık matrisinden gmean hesaplanması60

Denklem 3.27. Karmaşıklık matrisinden mcc değerinin hesaplanması.....61



SİMGELER VE KISALTMALAR DİZİNİ

6MWT	:6-Min Walk Test
ABC	:The Activity Specific Balance Confidence Scale
ADA	:AdaBoost
ALS	:Amyotrophic Lateral Sclerosis
ANN	:Artificial Neural Network
BDÖ	:Berg Denge Ölçeği
BPHDÖ	:Birleştirilmiş Parkinson Hastalığı Değerlendirme Ölçeği
CWT	:Comfortable 10-M Walk Test
DLB	:Dementia With Lewy Bodies
DMD	:Duchenne Musküler Distrofi
DVM	:Destek Vektör Makinaları
FUT	:Fonksiyonel Uzanma Testi
FWT	:Fast 10-M Walk Test
GN	:Gerçek Negatif
GP	:Gerçek Pozitif

GRF	:Ground Reaction Forces
k-EYK	:K En Yakın Komşu Algoritması
KNN	:K Nearest Neighbors Algorithms
LASSO	:Least Absolute Shrinkage And Selection Operator (En Küçük Mutlak Daralma ve Seçme Operatörü)
LDA	:Linear Discriminant Analysis
Leaveout	:Leave-one-out (Birini Dışarda Bırakma)
LR	:Lojistik Regresyon
MID	:Ortak Bilgi Farkı
MIQ	:Ortak Bilgi Oranı
miniBEST test:	Mini Balance Evaluations Systems Test
MR	:Manyetik Rezonans
MRG	:Manyetik Rezonans Görüntüleme
mRmR	:Minimum Redundacy Maximum Relevance (Minimum Artıklık Maximum İlişki)
MSE	:Ortalama Kare Hata
NB	:Naive Bayes
ÖFU	:Öne Fonksiyonel Uzanma
REM	:Rapid eye movement (hızlı göz hareketi)
RF	:Random Forest
RMS	:Root Mean Square (Karesel Ortalamanın Karekökü)

RO	:Rastgele Orman
SBS	:Ardışık Geri Yönde Seçim
SFS	:Ardışık İleri Yönde Seçim
SMMT	:Standardize Mini Mental Testi
SVM	:Support Vector Machines
TND	:Türk Nöroloji Derneği
TPDDYÖ	:Tinetti Performansa Dayalı Denge ve Yürüme Ölçeği
UPDRS	:Unified Parkinsons Disease Rating Scale
YFU	:Yana Fonksiyonel Uzanma
YN	:Yanlış Negatif
YP	:Yanlış Pozitif
YSA	:Yapay Sinir Ağları

BÖLÜM 1

GİRİŞ

Parkinson hastalığı, Alzheimer hastalığından sonra dünya üzerinde en sık görülen ikinci nörodejeneratif hastalıktır. Dünya üzerinde 10 milyondan fazla kişinin Parkinson hastalığından etkilendiği düşünülmektedir ve yalnızca Amerika Birleşik Devletlerinde 51,9 milyar dolarlık bir ekonomik etkiye sahip olduğu tahmin edilmektedir (Oliveira, Gasser, Edwards, Zweckstetter, Melki, Stefanis & Outeiro, 2021). Nörodejeneratif hastalıklar; nöronların yani sinir hücrelerinin yapılarının bozulmaları veya nöronların tamamen ölmeleri sonucu çeşitli belirtilerle ortaya çıkan hastalıklardır. Parkinson, Huntington, Alzheimer nörodejeneratif hastalıkların örnekleridir. Parkinson da dâhil olmak üzere, nörodejeneratif hastalıkların tamamı sinir hücrelerinin kendilerini yenileme özelliklerinin olmaması nedeniyle tedavisi mümkün olmayan hastalıklardır. Ancak bu tür hastalıklar erken teşhis edildiğinde hastanın yaşam kalitesi olumsuz etkilenmeden, ilerleyişi yavaşlatılabilir. Bu sebeple erken teşhis için geliştirilecek her yeni yöntem Parkinson hastalığı ile mücadele edilmesinde önemlidir.

Parkinson hastalığının sebebi olarak beynin Substantia Nigra olarak tanımlanan bölgesinde meydana gelen nöronal bozulma gösterilmektedir. Nöronların iyileştirilmesi mümkün olmadığından Parkinson hastalığının tamamıyla tedavisi de mümkün değildir.

Tedavi edilemeyen bu hastalıkta sadece uygun dozda ilaç ile hastalığın ilerlemesi durdurulabilir ya da yavaşlatılabilir ve belirtileri bastırılabilir. Bu hastalığın tedavisi anlamına gelmez yalnızca hastanın yaşam koşullarının iyileştirilmesi için gereklidir (Yiğit & Arıcıoğlu, 2015).

Parkinson hastalığının bir tedavisi olmadığı gibi kesin bir tanı veya test yöntemi de yoktur. Hastalığın sinsi ilerlemesi, teşhisini zorlaştırmaktadır. Ayrıca gösterdiği belirtiler incelendiğinde diğer nörodejeneratif hastalıklar ve demans gibi yaşlılarda sık görülen diğer hastalıklar ile benzer belirtiler gösterdiği gözlemlenmektedir. Bu durum hastalığın teşhisini zorlaştırmakta dolayısı ile müdahaleyi geciktirmektedir. Teşhis için bazı tahlillere ek olarak hasta ve yakınlarının beyanları dinlenmekte ve nörolog tarafından fiziki muayene yapılmaktadır. Kesin teşhis için görüntüleme yöntemlerinin (Manyetik Rezonans (MR) görüntüleme ve/veya Radyolojik görüntüleme gibi) tercih edilmesi de hem hastanın radyasyona maruz bırakılması hem de sağlık kuruluşunun maliyetli yöntemleri kullanmak zorunda kalması açısından dezavantajlıdır.

Parkinson hastalığında beynin Substantia Nigra bölgesindeki dejenerasyonun sebebi günümüzde halen tam olarak bilinmemektedir. Ancak hastalığın genetik yatkınlık ile bağıntılı olduğu düşünülmektedir. Parkinson hastalığına genetik yatkınlığı olan bireylerin hastalığı tetikleyici çevresel faktörlere de maruz kalması ile hastalığın geliştiği düşünülmektedir.

Parkinson hastalarında kas ve eklem katılıkları beraberinde yürüyüş ve denge sorunları yaşamalarına da neden olmaktadır. Özellikle ileri evrede Parkinson hastalarının yürüyüşlerinde belirgin farklılıklar gözlemlenmektedir. Kimi hastada belirtilerin tek taraflı görülmesi de yürüyüşün bozulmasında önemli bir etken olabilmektedir. Parkinson hastaları tipik olarak küçük adım aralıkları ile ve ayaklarını sürükleyerek yürümektedirler. Ayrıca hastalar denge problemi yaşadıkları için tipik olarak ağırlık merkezini yakalamak istercesine öne eğik olarak yürümektedirler. Hastalar yürüyüşü başlatma, yürüyüş esnasında sağa-sola dönme, durma, çömelme, oturma ya da kalkma hareketlerinde de zorlanmaktadır. Tüm bu bulgular Parkinson hastalığının teşhisinde yürüyüş parametrelerinin kullanılması fikrini ortaya çıkarmış ve başarılı sonuçlar elde edilmesini sağlamıştır. Literatürde bu konuda yapılmış farklı alanlardan araştırmacıların yaptığı çok

sayıda çalışma bulunmaktadır. Sensör teknolojilerinde yaşanan gelişmeler, objektif sonuçlar veren yeni sistemlerin keşfedilmesi, bilgisayar ve yazılım alanındaki gelişmeler ve tıp ve mühendislik alanında yapılmak istenen disiplinler arası biyomedikal çalışmalar bu konunun güncelliğini korumasını ve yeni araştırmacıların da bu konuya yönelmesini sağlamaktadır.

Tıp alanında çalışmalar yapan araştırmacılar tedavi yöntemlerinin geliştirilmesi ve başarısının analizi, fizik tedavi ve rehabilitasyon alanındaki araştırmacılar rehabilitasyon süreçlerinin objektif izlenebilmesi, eczacılık alanındaki araştırmacılar ilaçların farmakolojik etkileri, genetik bilimciler Parkinson hastalığına neden olan genlerin keşfi yönünde çalışmalar yaparken, mühendislik alanındaki araştırmacılar ise sensör ve yazılım teknolojisi yardımıyla nörologlara yardımcı yeni sistemler geliştirmeyi amaçlayan çalışmalar yapmaktadırlar. Mühendislik alanında yapılan bu çalışmalar nörologların daha objektif sonuçlar almalarını sağlamanın yanı sıra tanı süreçlerinin kısılmasını da sağlamaktadır. Aynı zamanda hastalara invaziv (girişimsel) hiçbir müdahale yapılmaması da bir avantajdır. Hastanın sensörler kullanılarak yürüyüş parametrelerinin kaydedilmesi hastanın radyasyon gibi zararlı etkilere maruz kalmasını da engellemektedir.

Yürüyüş tıpta nörolojik ve ortopedik alanlardaki hastalıkların teşhisinde ve tedavi sürecinin izlenmesinde sıkça kullanılan parametrelerden birisidir. Parkinson hastalığı, Serebral palsy, Duchenne Musküler Distrofi (DMD), Amyotrophic lateral sclerosis (ALS), Huntington hastalığı, Perthes hastalığı yürüyüş analizi ile incelenen hastalıklar arasında sayılabilir. Bu hastalıkların her birinin yürüyüş üzerine etkileri ayrı ayrı incelenmiş ve her bir hastalığın yürüyüşün farklı bir veya birden fazla parametresi üzerinde etkisi olduğu ispatlanmıştır.

Parkinson hastalığında yürüyüşün incelenmesi geçmiş yıllarda araştırmacıların sıkça incelediği konulardandır. Ancak konu halen güncelliğini korumaktadır. Özellikle giyilebilir sensör teknolojilerinde her geçen gün yaşanan gelişmelerle yürüyüşün tüm özellikleri ayrıntılı olarak çıkartılabilmekte ve incelenebilmektedir. Sayısal verilerin elde edilmesi, araştırma sonuçlarının objektifliğini arttırmakta ayrıca istatistiksel analiz, makine öğrenmesi ve yapay zekâ algoritmaları gibi farklı alanlarda çalıştırılarak detaylı sonuçlar elde edilmesine de imkân sağlamaktadır.

Bu çalışmada Malezya MARA Teknoloji Üniversitesi'nden Prof. Nooritawati Md Tahir ve ekibinin Parkinson hastası ve sağlıklı bireylerden elde ettiği yürüyüş verileri kullanılmıştır. Çalışmada hasta-sağlıklı ayrımı yapan makine öğrenmesi temelli bir algoritma geliştirilmesi amaçlanmıştır. Verilere farklı öznitelik çıkarma yöntemleri uygulanarak verilerin öznitelik sayıları artırılmıştır. Ham veride toplamda 21 olan öznitelik sayısı 153'e çıkarılmıştır. 153 öznitelik; Sıfır Ortalama ve Min-Maks Normalizasyon olmak üzere iki farklı normalizasyon yöntemine tabi tutulmuştur. Elde edilen veriler SVM (Support Vector Machines-Destek Vektör Makinaları (DVM)), KNN (k-nearest neighbors algorithm-K En Yakın Komşu (k-EYK)) Algoritması ve Naive Bayes Sınıflandırıcısı olmak üzere üç farklı sınıflandırma algoritmasında çalıştırılmıştır. Sınıflandırıcıların performanslarını test etmek amacıyla "Birini Dışarda Bırak (Leave-one-out)" ve "Kfold" çapraz doğrulama yöntemleri kullanılmıştır. Son olarak da Lasso (Least Absolute Shrinkage and Selection Operator- En Küçük Mutlak Daralma ve Seçme Operatörü), Relief ve mRmR (Minimum redundancy maximum relevance-Minimum Artıklık Maksimum İlişki) öznitelik seçme yöntemleri uygulanarak en belirleyici öznitelikler elde edilerek, sonuçlar karşılaştırılmış ve yorumlanmıştır.

Bu tez çalışmasında Bölüm 2, çalışmanın ana konusu olan Parkinson hastalığı, Parkinson hastalığının yürüyüş ve denge üzerindeki etkisi ve belirtileri ile ilgili genel tanımları ve yorumları içermektedir. Bölüm 3'te kullanılan makine öğrenmesi ve öznitelik çıkarım yöntemleri hakkında bilgilendirme yapılmıştır. Bölüm 4'te yapılan çalışmalar, elde edilen sonuçlar ve karşılaştırmaları ayrıntılı olarak verilmiştir. Sonuçların ayrıntılı ve karşılaştırmalı yorumlanması, tartışma ve ileride yapılacak çalışmalar hakkındaki bilgilere Bölüm 5'te yer verilmiştir.

BÖLÜM 2

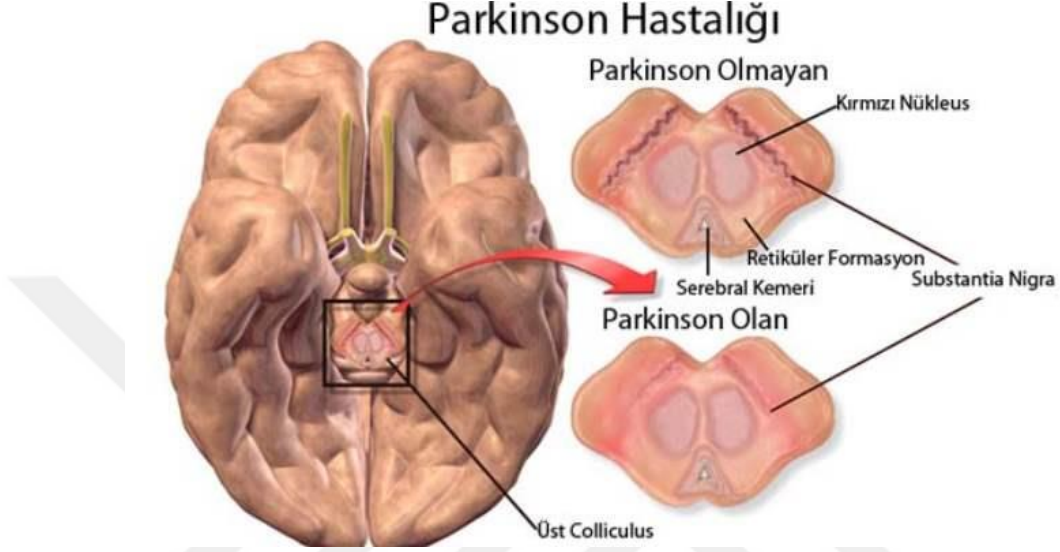
PARKİNSON HASTALIĞI VE YÜRÜYÜŞE ETKİSİ

2.1. Parkinson Hastalığı Hakkında Genel Bilgiler

Parkinson hastalığı, ilk kez 1817 yılında İngiliz hekim ve bilim insanı James Parkinson tarafından “Titremeli Felç Üzerine Bir Deneme (An Essay on the Shaking Palsy)” isimli kitabında tanımlanmıştır. Parkinson hastalığı dünya üzerinde Alzheimer hastalığından sonra en sık rastlanan ikinci nörodejeneratif hastalıktır. Hastalığın en yaygın belirtileri motor semptomlarda gözlenen titreme (tremor), katılık (rijidite) ve hareketlerde yavaşlama (bradikinezi) şeklindedir. Ayrıca bilişsel, davranışsal ve otonomik şikâyetler de Parkinson hastalığının belirtileri arasında sıklıkla görülmektedir (Baydın, 2021).

Parkinson hastalığı; beynin Substantia Nigra olarak isimlendirilen bölgesinde bulunan nöronların ölmesi sebebiyle ortaya çıkan bir hastalıktır (Kasap & Akpınar, 2011). Substantia Nigra, orta beyin bölgesinde bulunan, dopamin üreten ve harekette önemli rol oynayan bir yapıdır. Substantia Nigra’da meydana gelen dejenerasyon ile birlikte dopamin üretimi gerçekleşemez veya sınırlı olarak gerçekleşir. Dopamin ise hareketlerin akıcılığını ve uyumluluğunu sağlayan, sinir hücreleri arası haberleşmeye katkısı olan bir maddedir. Substantia Nigra’da meydana gelen dejenerasyon dopamin üretimini sınırlandırır, bu da

istemsize titreme (tremor) hareketlerine, hareketlerde yavaşlamaya (bradikinezi), katılığa (rijidite), denge sorunlarına, duruş bozukluđuna (postral instabilite) ve yürüme bozukluđu gibi bir çok bulgunun gelişmesine yol açar (Altun, Özbek, Zarifođlu & Özkaya, 2013).



Şekil 2.1. Parkinson hastası olan ve olmayan bireylerin beyinlerindeki Substantia nigra farklılıkları (Hemensaglik, 2021).

Hastalığın sebebinin Substantia Nigra bölgesindeki nöronların işlevlerini tam olarak yerine getirememesi olarak bilinmektedir ancak bu bölgedeki dejenerasyonun nedeni hakkında kesin bir bilgi bulunmamakla birlikte genel kanı hastalığın genetik kökenli olduđu yönündedir. Genetik yatkınlık, çevresel tetikleyici faktörler ile de birleşince hastalığın ortaya çıktığı düşünölmektedir. Buradaki çevresel faktörler 2002 yılında yapılan bir çalışmada toksik maddelere maruz kalma, sigara kullanımı, kullanılan ilaçlardan kaynaklanan yan etkiler vb. şeklinde açıklanmaktadır (Di Monte, Lavasani & Manning-Bog, 2002).

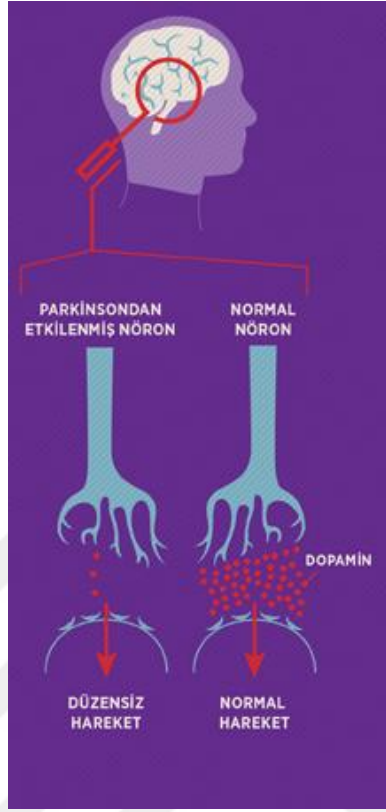
Yaşlı bireylerde Parkinson hastalığının görülme sıklığı genç bireylere göre daha fazladır. Örneğin; 50-59 yaşları arasındaki bireylerde görülme sıklığı 100.000'de 17,4 iken; 70-79 yaşları arasında görülme sıklığı ise 93,1'dir. Hastalığın ortalama başlangıç yaşı 60-80 aralığında olarak ifade edilmektedir. Parkinson hastalığı daha genç yaşlarda da görülebilir; 40 yaşından önce Parkinson hastalığının başladığı vakalar "Genç Başlangıçlı Parkinson" olarak nitelendirilirken, 20 yaşın altında görölen vakalar ise "Jüvenil Parkinson" olarak tanımlanmaktadır (Yaliman & Şen, 2011).

Parkinson hastalığında cinsiyet ve yaş, hastalığın oluşumu ve seyrinde önemli faktörlerdir. Özellikle gelişmiş ülkelerde ortalama ömrün uzaması ile Parkinson hastalarının sayısı gün geçtikçe artmaktadır. Türkiye’de Parkinson hastası birey sayısının 120.000 civarında olduğu düşünülmektedir. Ayrıca, 50 yaşını geçmiş erkeklerde Parkinson hastalığının görülme sıklığı kadınlara göre daha yüksektir (Kurt, 2019).

Parkinson hastalığının tanısında UPDRS (Unified Parkinsons Disease Rating Scale) - (Birleşik Parkinson Hastalığı Değerlendirme Ölçeğiyle (BPHDÖ)) olarak isimlendirilen bir ölçek kullanılmaktadır. UPDRS ölçeği günümüzde en sık kullanılan ve en geçerli yöntemdir. UPDRS ölçeği depresyon, anksiyete (duygu durumunda kaygı, endişe artışı), apati, uyku sorunları gibi nonmotor sorunların yanı sıra konuşma, yeme, yazma, yatakta dönme, titreme, yürüme ve denge gibi motor sorunları da içeren bir ankettir. Hasta ve yakınlarının beyanına dayanan bu anket bölümlerinden sonra hekimin motor muayene bulgularını kaydettiği ve konuşma, yüz ifadesi, rijidite, parmak hareketi, el hareketi, postural denge ve yürüme-donma bölümlerinden oluşan puana dayalı ölçek bulunmaktadır. Ölçeğin son bölümünde de motor komplikasyonların değerlendirilmesi yer almaktadır. UPDRS ölçeklendirmesi tamamen hasta ve hasta yakını beyanına ve hekimin gözlemine göre hastanın o anki durumunun puanlandırılmasına dayanan bir ölçektir. Bu puanlandırma nicel olmayan az-çok gibi derecelendirmeler içermektedir.

2.1.1. Parkinson Hastalığının Belirtileri

Parkinson hastalığının belirtileri arasına donuk yüz ifadesi, göz kırpma sayısında azalma, mimiklerde azalma ya da katılaşıma, kambur duruş, ses tonunda değişme, konuşma bozukluğu, ağızda salya birikmesi, yutmada zorlanma, kabızlık, idrar yapma kusurları, cinsel işlev bozuklukları, aşırı terleme, giderek küçülen ve titremeye bağlı olarak karmaşıklaşan el yazısı, koku almada bozulma, çeşitli ağrılar, depresyon, unutkanlık, uyku bozukluğu vb. eklenebilir. Belirtilerin tamamının bir hastada görülmesi şart değildir ancak bahsi geçen belirtilerden birkaçının görülmesi dahi hastanın yaşam kalitesini doğrudan olumsuz yönde etkilemektedir. Aynı zamanda hastalığın ileri evrelerde olması durumunda bu belirtiler daha da ilerleyerek hastanın hayatını tek başına idame ettirmesine engel olmaktadır. Öyle ki hastalar yemek yeme, kaşık tutma, yürüme, giyinme gibi günlük basit hareketleri bile tek başlarına yapamaz hale gelebilirler.



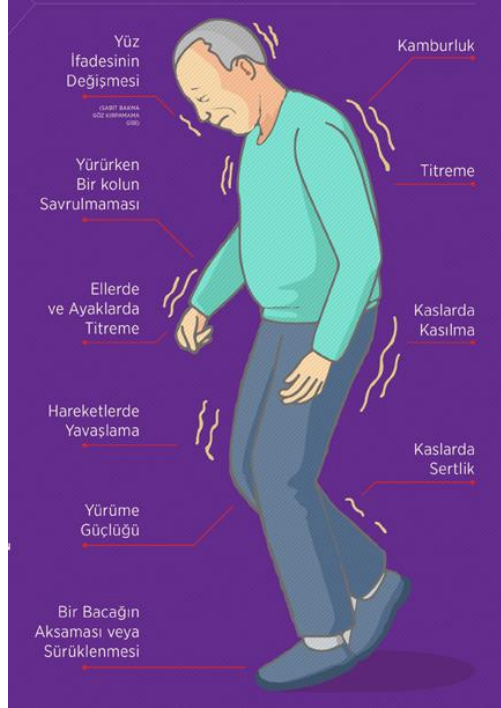
Şekil 2.2. Dopamin'in Parkinson hastalığına etkisi (Gözellik, 2018).

- **Titreme (Tremor):** Titreme, Parkinson hastalarının en belirgin hissettiği ve en çok şikâyetle doktora başvurduğu belirtidir. Parkinson hastalığında titremenin en temel özelliği dinlenme halinde görülmesidir. Bu sebeple “istirahat tremoru” olarak adlandırılmaktadır. Hasta popülasyonunun ortalama %70’inde titreme görülmektedir. Parkinson hastalarında titreme frekansı 4-6 Hz iken, essential titremede 5-8 Hz, fizyolojik titremede ise 8-12 Hz aralığındadır. El veya ayakta görülen titreme zamanla ilerleyerek kol ve tüm bacak bölgelerine yayılabilir. Parkinson hastalığında dil, çene ve dudaklarda da titreme görülebilmektedir.
- **Katılık (rijidite):** Katılık, ekstremitelerin (uzuvların) kaslarının sertliği olarak da isimlendirilmektedir. Katılık hasta tarafından fark edilmesi zor bir histir. Çoğunlukla muayene esnasında hekim tarafından elle muayene edilirken fark edilir. Hastalığın ileri evrelerinde omurga çevresi kaslara kadar sertlik ilerleyebilir.
- **Hareketlerde Yavaşlama (Bradikinezi):** Her hastada farklı şiddette görülmekle birlikte, hareket yavaşlığı çoğunlukla yaştan kaynaklı olabileceği düşünüldüğünden

doktora geç başvurulmasına da neden olan bir belirtidir. Hastalar gündelik hayattaki hareketlerinde örneğin; kravat bağlama ve takma, düğme ilikleme, çatal-kaşık kullanma ve yazı yazma gibi hareketlerde zorlanma yaşar. Zamanla bu belirti de ilerleyerek gündelik hayattaki birçok aktiviteyi yapamamaya neden olabilmektedir. Örneğin; yatar pozisyonda iken dönme hareketini gerçekleştirememesi, arabaya binememesi ya da inememesi, yemek yeme esnasında çiğneyememesi, koltuğa oturma, koltukta doğrulma ve koltuktan kalkma hareketlerini gerçekleştirirken önemli derecede zorlanırlar.

- **Yüz ifadesinde donukluk (hipomimi):** Parkinson hastalarında yüz ifadesini sağlayan yüz kaslarının katılaşması ve yüz hareketlerinin azalması ile donuk bir yüz ifadesi görülür. Ayrıca hastalarda otomatik göz kırpması hareketinde azalma görülmektedir. Hasta çoğu zaman bu durumun farkında olmayabilir, ancak hasta yakınları tarafından dikkatini çeker.
- **Konuşma ve Ses Bozuklukları:** Parkinson hastalarında ses üretmede ve konuşmada bozulma, dopamin eksikliğinin yüz, ağız gırtlak ve yutak kaslarının sertleşmesi ve hareket kabiliyetinin azalması nedeniyle görülmektedir.
- **Mikografi:** Hastalar özellikle imza atma esnasında zorlanma yaşadıklarında belirtinin farkına varırlar. Bu durum titreme ve hareket yavaşlığı ile de bağdaştırılabilir. Hastalarda el yazısı yazarken giderek küçülen ve kimi zaman karmaşıklaşan el yazısı gözlemlenmektedir.
- **Ruhsal, Zihinsel, Duysal, Duyusal Bozukluklar:** Ruhsal çöküntü hali, depresyon, unutkanlık, demans, ağrılar, duysal yakınmalar, koku duyusu kaybı bu alt başlığa örnek gösterilebilir. Hastaların önceki hallerine göre durgunluk, neşesizlik, iştahsızlık duygularındaki artış sık görülen belirtilerdir. Hastalığın ileri dönemlerinde hafif veya orta derecede unutkanlık yaşa da bağlı olarak görülebilmektedir. Hastalığın sık görülen belirtilerinden olan kas sertliği beraberinde kas ağrılarını da getirebilmektedir. Ancak bunun dışında hastalığın vücudun tek tarafında görülmesi de ağrıların tek taraflı olarak hissedilmesine neden olabilmektedir. Hastaların ilaç tedavisine başlamaları ya da ilaç tedavisi görmekte olan hastanın ilaç dozunun düzenlenmesi durumunda ağrıların da paralel olarak azaldığı ve/veya tamamen ortadan kalktığı görülmektedir. Nadiren hastaların el ve ayaklarda üşüme hissinden yakındıkları da görülür

- **Uyku Bozuklukları:** REM (Rapid eye movement) uykusu evresinde hastaların sağlıklı kişilere kıyasla daha hareketli olduğu saptanmıştır. REM evresi; uykuda hızlı göz hareketlerinin olduğu ve rüya görülen evredir. Aynı zamanda uykuda EEG kaydı alındığında kişiler diğer uyku evrelerinden farklı olarak uyanıklıktaki beyin sinyallerine benzeyen sinyaller üretirler. Özellikle erkek hastaların uyku esnasında konuşma, bağırma, yumruk ya da tekme atma hareketleri görülebilir. Ayrıca Parkinson hastaları gündüzleri uyuklama halinden de şikâyetçidirler.
- **Duruş ve Yürüyüş Bozuklukları:** Özellikle hastalığın tek taraflı olarak başladığı bireylerde yürüyüş esnasında bir bacağın diğerine göre farklı davranış sergilemesi, yürüyüşteki bozulmanın en belirgin göstergesidir. Parkinson hastalığının yürüyüşe etkileri son derece değerlidir. Hastalığın yürüyüşe başlıca etkileri; yürüyüşte yavaşlama, ayak sürükleme, yürürken ağırlık merkezini dengelemek için öne doğru eğik ve küçük adımlarla yürüme, yürüyüşte düzensiz kol salınımı, ayak tabanı ağırlık noktalarının öne doğru kayması, yürürken kilitlenme (donma), donmaya bağlı ani düşmeler, yürüyüş esnasında topuk açısı, kalça açısı ve diz açısında farklılaşmalar vb. şeklindedir.



Şekil 2.3. Parkinson hastalığının bazı belirtileri (Gözellik, 2018).

Parkinson hastalığının tedavisinde başvurulan en genel yöntem ilaç tedavisidir. İlaçlar yetersiz dopamin maddesinin açığını kapatarak hasta şikâyetlerinin azalmasını sağlamaktadır. İlaçlar belirtileri baskılayarak hastanın günlük hayat konforuna katkıda bulunur. Ancak ilaçlar hastalığın tedavisini sağlamamaktadır. İlaç kullanılmasına rağmen hastalık ilerleyebilir ve alınan ilacın dozu yetersiz kalabilir. Bu sebeple sık kontrollerle hastalığın ilerleyişi hekim izlemi gerektirir.

Ek olarak, belirtilerin baskılanması için beyin pili implantasyonu da bir seçenek olarak son yıllarda sıklıkla kullanılmaktadır. Beyin pili uygulamaları hastalığın ilerleyen evrelerinde nörologların uyguladığı bir yöntemdir. Çünkü hastalığın ilk evrelerinde belirtiler baskın olmaz ve ilaç tedavisinden olumlu yanıt alınır. Derin beyin stimülasyonu olarak da isimlendirilen bu cerrahi yöntem, özellikle istemsiz hareket ve titreme gibi motor bulguların baskılanmasında başarılı sonuçlar vermektedir (Yiğit & Arıcıoğlu, 2015).

2.2. Yürüyüş ve Denge

Yürüme; bir yerden farklı bir yere gidebilmek amacıyla tekrarlanan ritmik hareketler olarak tanımlanabilir. İlgüdülere dayanan temel yürüyüş döngüsü üzerine öğrenilen davranışların da katılması ile ırklara, bireye hatta bireyin o anki ruh haline özgü yürüyüş davranışı gelişebilir. Yürüyüş hareketinin analizi M.Ö. 320 yıllarına kadar dayanmaktadır. Ancak günümüzde algıladığımız anlamda üç boyutlu olarak yürüme analizini klinik kullanıma sokan araştırmacılar Verne Inman ve Jacquelin Perry olmuştur (Yavuzer, 2014).

Uzun müddet yürüme fonksiyonunu devam ettirebilmek için beyin, omurilik, sinirler, kemik, kaslar ve eklemler senkronize olarak çalışmalıdır. Eklem hareketleri sağlıklı olarak gerçekleştirilebilmeli ve kas kasılmalarının gerçekleşebilmesi için yeterli güç olmalıdır. Vücut bölümlerinin hareketleri ve kas kasılma-gevşeme zamanlamaları yürüme için gerekli olan enerji ihtiyacını minimum düzeyde olacak şekilde tasarlanmıştır. Bu uyumu aksatan her türlü yapısal bozukluk ya da fonksiyon bozukluğu enerji ihtiyacını arttırır ve yürüyüş verimini düşürür (Yavuzer, 2014).

İnsan vücudunda denge, karmaşık nöromüsküler sistem (kas ve sinir sistemi) tarafından değişen koşullara karşı sürekli bir uyum halinde olması olarak açıklanabilir. Dinlenme ve aktivite esnasında, yerçekimi merkezinin değişikliklerine karşı vücut ağırlık merkezinde de hızlı bir postüral uyum gerçekleşir. İşte bu uyuma denge denir (Soyuer & İsmailoğulları, 2009).

Günümüzde birçok hastalığın teşhisinde ve tedavisinin izlenmesinde yürüyüş ve denge testlerinden faydalanılmaktadır. Ortopedik ve nörolojik hastalıkların yürüyüş üzerine olan etkisinin incelenmesi geçmişten bu yana halen güncelliğini koruyan yöntemlerdendir. Parkinson hastalığı, Alzheimer hastalığı, inme, Serebral palsy, Duchenne Musküler Distrofi (DMD), Amyotrophic lateral sclerosis (ALS), Huntington hastalığı, omurilik hasarları, Perthes hastalığı yürüyüşün incelendiği hastalıklar arasında sayılabilir. Ayrıca ampute bireylerde protez modellerinin karşılaştırılmasında, spor yaralanmalarında, spastisite tedavisinde yürüyüş analiz yöntemleri kullanılmaktadır. Ek olarak ortopedide implantların, ortezlerin ve protezlerin yürüyüşe etkilerinin incelendiği çalışmalar da bulunmaktadır.

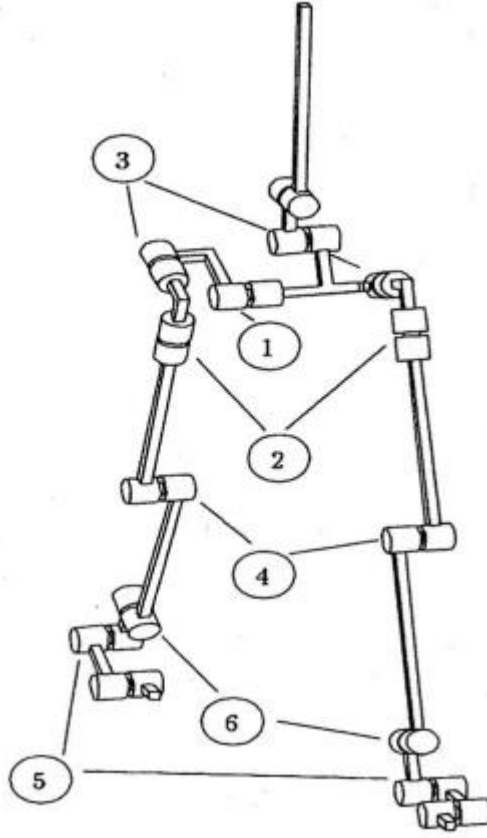
Yürüme analizi, kas ve sinir sistemi fonksiyonlarının sayısal olarak elde edilmesi ve sonuçların sayılar ve grafikler ile yorumlanması şeklinde tanımlanabilir.

Yürüyüşte incelenen bazı özneliklere yürüyüş hızı, kadans (dakikadaki adım sayısı), tek adım uzunluğu, çift adım uzunluğu, adım genişliği, ayak açısı, diz açısı, kalça açısı, destek yüzeyi, tek bacak duruş süresi örnek verilebilir.

Yürüyüş ve dengenin değerlendirilmesinde literatürde yer alan farklı ulusal ve uluslararası test teknikleri bulunmaktadır. Bunlar: Standardize mini mental testi (SMMT), Fonksiyonel Uzanma Testi (FUT), Öne fonksiyonel uzama (ÖFU), Yana fonksiyonel uzanma (YFU), Tinetti performans dayalı denge ve yürüme ölçeği (TPDDYÖ), Berg denge ölçeği (BDÖ), Türk Nöroloji Derneği (TND) denge ve yürüme bozuklukları değerlendirme formu, Tek ayak üzerinde durma, 10 m yürüme testi, Sandalyeye otur-kalk testi, Süreli Kalk Yürü Testi, Zamanlı kalk yürü testi şeklinde sıralanabilir.

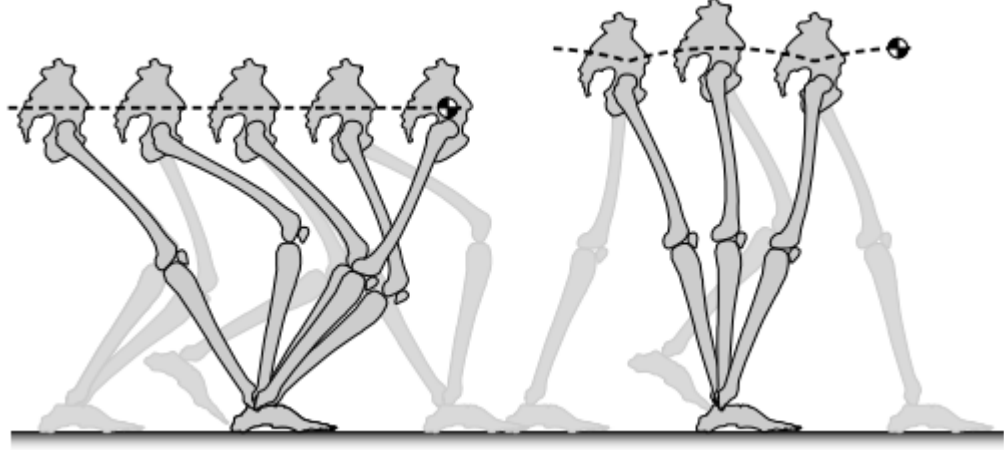
2.2.1. Yürüyüş Anatomisi

Yürüyüşün temel özellikleri ilk olarak 6 madde halinde Saunders vd. Tarafından 1953 yılında tanımlanmıştır. Bu altı maddelik temel yürüyüş belirleyiciler: pelvik rotasyonu, pelvik eğim, duruş diz fleksiyonu, tibiofemoral açı, ayak ve ayak bileği mekanizmalarıdır (Hayot, Sakka & Lacouture, 2011). Saunders vd. (1953)'nin ortaya attığı 6 belirleyici Şekil 2.4'de gösterilmiştir.



Şekil 2.4. Saunders ve ark. (1953) tarafından tanımlanan 6 temel yürüyüş belirleyici (de Pina Filho, Dutra & Raptopoulos, 2005).

Bu altı belirleyici esasen yürüyüş esnasında, bir dizi kinematik özniteliğin vücut ağırlık merkezinin yer değiştirmesini azaltmaya yardımcı olduğunu öne sürer. Ancak bu teoriye karşıt bir teori de 1966 yılında Cavagna vd. tarafından literatüre eklenmiştir. Cavagna vd.'nin teorilerine göre ise Saunders vd.'nin aksine yürüme esnasında vücut ağırlık merkezi sabit kalmaz, bir sarkaç gibi hareket eder (Cavagna & Margaria, 1966; Kuo, 2007). Her iki teoriye ait temsili bir görsel Şekil 2.5'te gösterilmiştir.

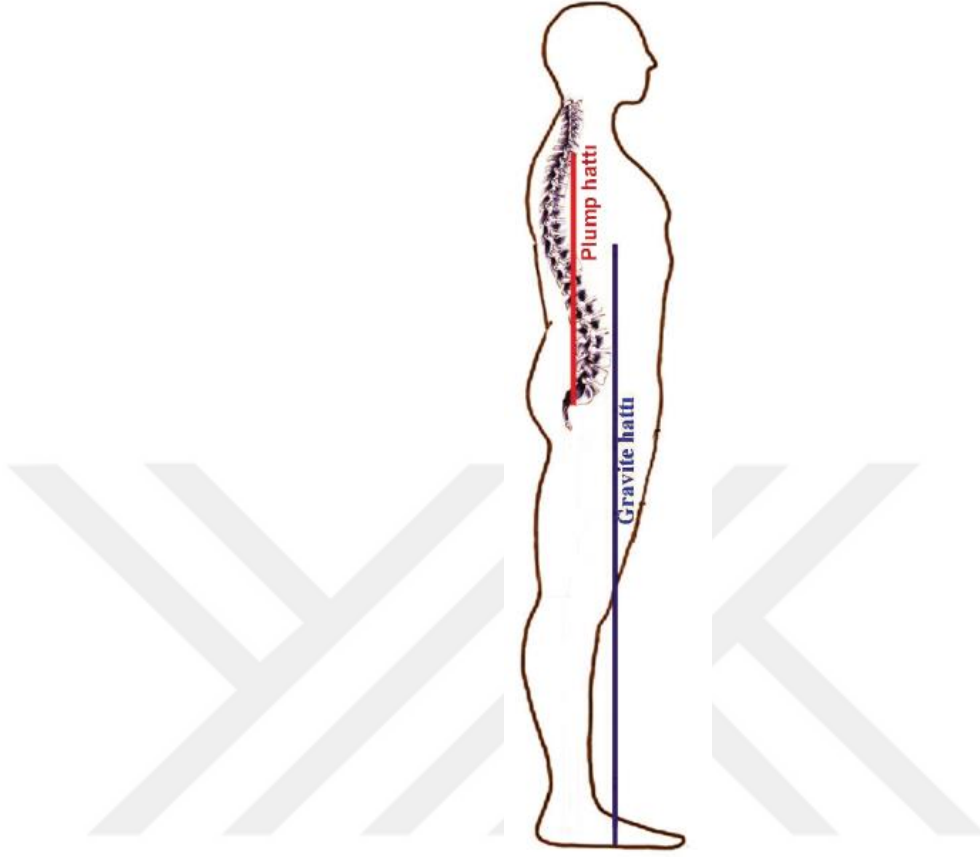


Şekil 2.5. Yürüyüş döngüsünde ağırlık merkezinin hareketi ile ilgili iki farklı teoriye (Saunders vd., 1953 ve Cavagna vd., 1966) ait görsel (Kuo, 2007).

Akıcı ve fazla enerji ihtiyacı doğurmayan yürüyüş normal yürüyüş olarak değerlendirilir. Ayakta duran bir insanın vücut ağırlık (gravite) merkezi ikinci sakral omurun önünde bulunur (Şekil 2.6). Yürüyüş sırasında ağırlık merkezi sinüzoidal bir eğri çizer. Yürüme esnasında her adımda pelvis yukarı ve aşağı yönlü toplamda 9-10 cm hareket eder ve 4° sağa - 4° sola dönme hareketi gerçekleştirir. Birçok hastalıkta vücut ağırlık merkezinin yukarı-aşağı, sağa-sola hareketinde değişiklikler gözlenir (Alsancak, 2015).

Sağlıklı bireylerde yürüyüş hızı ortalama 80 m/dk'dır. Bireylerin yürüyüş hızındaki değişiklikler de bazı hastalıkların belirtisi olabilir.

Yürüyüş hızı (m/sn) = adım uzunluğu(m) x kadans (adım/dk)/120 Denklem 2.1.

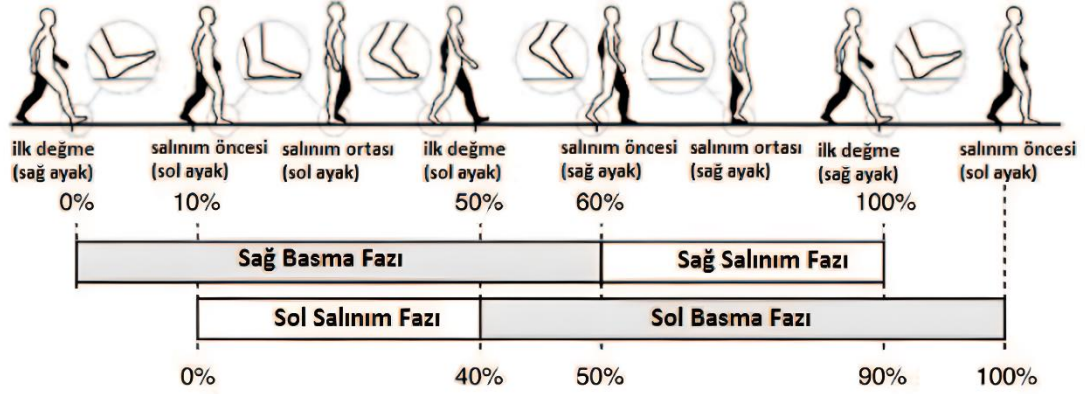


Şekil 2.6. Gravite (vücut ağırlık merkezi) hattı (Özer, Kaner & Bozdoğan, 2014).

Son dönemlerde yapılan çalışmalar incelendiğinde deneklerin belirli bir alanda, belirlenen miktarda yürütülmesi ile yürüme kayıtları alınmaktadır. Bu kayıtlar kamera tabanlı sistemler ya da giyilebilir sensörler içeren sistemlerden elde edilmektedir. Son dönemlerde yaşanan teknolojik gelişmelerle birlikte veri toplamada kullanılan bu sistemler gün geçtikçe daha da gelişmekte ve ulaşılabilir hale gelmektedir.

Yürüyüş analiz sistemlerinden elde edilen veriler çoğunlukla kinetik, kinematik ve temel yürüyüş verileri olarak sınıflandırılmaktadır. Kinematik veriler kısaca yürüyüş esnasında vücudun uzaydaki hareketi olarak tanımlanabilir. Yürüyüşte gövde, pelvis, bacak ve ayaklardaki eklemlerde oluşan açılar, hızlar ve ivmelerin sayısal verileri “kinematik veri” olarak adlandırılır. Kinetik veriler; ayak tabanındaki ağırlıkla oluşan yerin tepki kuvvetleri, eklem momentleri ve eklemlerin güçleri gibi kuvvet tabanlı verileri kapsar. Kinetik veriler özellikle fizik tedavi alanında kuvvet platformları ile

ölçülmektedir (Yavuzer, 2014). Temel yürüyüş verileri ise; yürüyüş hızı, adım süresi, kadans, adım uzunluğu vb. gibi hız ve uzunluk gibi değerleri kapsayan verilerdir.



Şekil 2.7. Yürüme döngüsü ve fazları (Pehlivan, Tunca, Salur & Ersoy, 2017).

Şekil 2.7’de yürüyüşte bir adımda gerçekleşen fazlar ve ayak, ayak bileğinde meydana gelen değişiklikler gösterilmiştir. Yürüyüş analiz sistemlerinde hastalar belirli bir mesafede yürütülür ve her bir adımları için Şekil 2.7’deki evrelerin açı, hız, ivme, konum vb. değerlerinin ortalaması alınır. Sağlıklı bir bireyde yürüyüşte bir adımın %60’ını basma fazı, %40’ını ise salınım fazı oluşturmaktadır. Bu yüzdeler dilimdeki değişiklik de hastalıkların teşhisinde kullanılabilir bir bulguya işaret eder. Şekil 2.7’deki görseli ele alacak olursak; basma fazı sağ ayak topuğunun yere temas etmesi ile başlar, sağ ayakta ağırlığın topuktan orta ve ön ayak (parmak) bölgesine doğru bölünmesi ve sol ayağın salınımı ile devam eder ve sol ayak topuğunun yere temas etmesiyle sağ ayağın yer ile temasının tamamen bitmesi ile sonlanır. Basma fazı esnasında bahsi geçen durumlarda vücut ağırlık merkezinin konumu ve hızı ve kalça açısı önemli ayırt edici bilgiler verir. Salınım fazı ise; basma fazının bitmesini takiben sağ ayağın tamamen yerden ayrılması ile başlar ve sağ ayağın topuğunun tekrar yere basması ile son bulur. Salınım fazında sağ ayak salınım halindedir. Salınım fazı da kendi içinde hızlanma, salınım fazı ortası ve salınım fazı sonu olmak üzere üç evreye ayrılabilir. Bu evrelerde ayak bileğinin açısı ve kalça ve dizdeki kasılma anlamlı bilgiler vermektedir. Bu evrelerden bağımsız olarak sağ ve sol ayağın aynı anda yer ile temasta olduğu süreler “çift destek faz”ları olarak adlandırılmaktadır. Sağ veya sol ayaktan yalnızca birinin yer ile teması hali ise “tek basma faz”ları olarak adlandırılır. Şekil 2.7’de bahsi geçen yürüyüş

evrelerinde oluşan topuk, diz ve kalça açıları, basma ve salınım faz süreleri hastalıkların teşhisinde anlamlı bilgiler vermektedir.

2.3. Yürüyüş, Denge ve Parkinson Hastalığı

Hareketi kontrol eden dopaminerjik sistemin merkezinde meydana gelen bozulmalar nedeniyle Parkinson hastalarında yürüyüş sırasında donma, yürüyüş yavaşlığı, adım uzunluğunda kısalma, küçük ve hızlı adımlarla yürüme ve ritim bozukluğu gözlenir. Bu belirtilerin hepsi bir hastada görülmeyebileceği gibi hastalığın ilerleyen evrelerinde belirtiler daha da hissedilir olur. Hastalığın ilk evrelerinde hareket sistemi ile ilgili belirtiler daha az hissedilmektedir. Ancak yine de belirtiler hastadan hastaya farklılık gösterebilir. Başlangıç evrelerinde hareket sistemi ile ilgili belirtiler uygun ilaç dozu ile baskılanabilmekte ve hasta konforu sağlanabilmektedir. Ancak ilerleyen evrelerde yüksek dozda ilaçlarla dahi hastalığın etkileri engellenemez. İşte bu sebeple hastalığın erken tespiti, hastalığın ilerlemesinin engellenmesi ve belirtilerin baskılanması için büyük önem taşımaktadır.

Parkinson hastalığının yürüyüş üzerindeki etkiler aşağıda maddeler halinde ve ayrıntılı olarak sıralanmıştır:

- Adım uzunluğunun kısalması,
- Ayak sürüklenme,
- Kadans değerinde anlamlı değişkenlik,
- Yürümeyi başlatmada güçlük,
- Denge problemleri,
- Yürürken donakalma,
- Yürüme esnasında soldan sağa veya sağdan sola dönme hareketinde güçlük ve/veya yavaşlama,
- Yürürken düzensiz kol salınım hareketi,
- Adım atmada salınım fazı içinde topuk, kalça ve diz açılarında azalma,
- Çift destek süresinde uzama,
- Yürümenin giderek hızlanması (kontROLSÜZ biçimde) (Altun vd., 2013).

Yukarıda maddeler halinde verilmiş olan hastalık etkileri beyinde salgılanması gerekenden daha az miktarda dopamin salgılanmasının sonuçlarıdır. Dopaminin vücutta eksik olması beraberinde kasların kontrol mekanizmasının da olumsuz etkilenmesini getirmektedir. Dolayısıyla kaslarını istemli kontrol edemeyen hasta, yürüyüşünü de kontrol edememektedir.

Parkinson hastalığında, kasların kontrol mekanizmasının etkilenmesi ile birlikte postural dengesizlik en sık görülen bir diğer belirtidir. Hastalar tipik olarak gövdeleri öne doğru eğilmiş bir görünüm sergiler. Bu durum hastanın vücut ağırlık merkezini yakalamak istemesine öne doğru eğik durmasını, küçük ve hızlı adımlarla yürümesini açıklamaktadır. Postural dengenin bozulması hastalarda ayağa kalkma, yürüme ya da duruş esnasında düşmeleri de beraberinde getirmektedir. Düşme ile meydana gelebilecek ciddi yaralanmalar özellikle ileri yaşlardaki hastalar için ciddi tehdit oluşturabilmektedir. Hastalığın ortalama yaşının yüksek olması düşmenin beraberinde getireceği yaralanmalar ile de farklı problemler yaşamalarına ve iyileşme sürelerinde uzamaya neden olmaktadır. Postural denge, Parkinson hastalığı tanısında hastalığın ilerleyişinin değerlendirilmesinde ve derecelendirilmesinde kullanılan en güvenilir ve geçerli ölçek olan UPDRS testlerinde de bulunan bir parametredir. Ancak UPDRS ölçeklendirmesinin motor bulgular bölümünde postural denge ile ilgili yalnızca bir maddelik bir test bulunmaktadır. Bahsi geçen UPDRS testleri ve puanlandırma tamamen hekimin takdirine ve gözlemine bağlı nicel olmayan bir ölçme işidir. Bu nedenle postural dengenin daha objektif ve nicel ölçümüne ihtiyaç duyulmaktadır.

Parkinson hastalarının yaşadığı postural dengesizlik ve yürüme sorunları hastaların bağımsız hareketini engellediği gibi aynı zamanda yaşam kalitesini de olumsuz etkilemektedir. Bu faktörlerin beraberinde getirdiği düşme ve düşmeyle birlikte yaralanma riski hastaların konforunu ve duygularını da olumsuz etkilemektedir. Hastalar hareket etmektan kaçınmaya başlar. Oysa literatürde hastalığın belirtilerinin baskılanmasında egzersiz ve hareket etmenin olumlu katkılarına olduğu ispatlayan çalışmalar vardır (Göz, Çolakoğlu, Çakmur & Balcı, 2020).

Parkinson hastalarının yürüyüşlerinde yalnızca yürüme hızı, dakikadaki adım sayısı (kadans) ve postral denge ile ilgili kayıtlar değerlendirilmemiştir. Bunlara ek olarak ayak tabanının farklı noktalarından kuvvet değerleri de değerlendirilerek sağlıklı kişilerin ayak kuvvet haritası ile olan farklılıkları da incelenmektedir. Sağlıklı kişilerden farklı olarak Parkinson hastalarının ayak tabanlarındaki kuvvet öne doğru ağırlıklanmış durumdadır. Bu durum hastaların postral dengesizliklerinin, duruş bozukluklarının ve öne doğru eğik duruş ve/veya yürüyüş göstermelerinin de bir sonucu olabilir. Özellikle hastalığın ileri evrelerinde topuk ve orta ayak bölgesinde ölçülen kuvvetin azaldığı, basıncın parmak uçlarında arttığı gözlemlenmektedir. Bu durum hastaların tipik olarak adım atma esnasında önce topuğu değil ayak parmak uçlarını yere temas ettirmelerini de açıklamaktadır. Bu sebeple Parkinson hastalarının yürüyüşlerinin incelenmesinde kuvvet sensörlü ayak tabanlıkları kullanılarak yapılan araştırmaların da sıklaşmasını sağlamıştır (Kimmeskamp & Hennig, 2001).

Yukarıda yer alan testlerin kimi sensör teknolojisi kullanılarak, kimi giyilebilir cihazlar kullanılarak kimi de fizik tedavi uzmanları tarafından gözlenip, puanlandırma yapılarak değerlendirilmektedir. Ancak elbette testler arasında sensörler vasıtası ile toplanan verilerin elde edilme yöntemi ve analizi açısından objektif olmaları ve kesin sonuçlar vermeleri bakımından daha avantajlı oldukları söylenebilir. Sensör teknolojisinin kullanımı, toplanan verilerin mühendislik yöntemleri ile işlenmesi ve yorumlanmasını kolaylaştırmaktadır. Ayrıca sensör kullanımının objektif değerlendirme, kullanımı kolay ve taşınabilir olma, uygun fiyat gibi diğer avantajları da bulunmaktadır (Hubble, Naughton, Silburn, & Cole, 2015).



Şekil 2.8. Yürüyüşün incelenmesinde kullanılan ve kablosuz veri aktarımı sağlayan bir kayıt cihazı örneği (Hasomed, 2021).

Günümüzde sensör tabanlı sistemlere ek olarak, 3 boyutlu hareket analiz sistemleri, kamera tabanlı hareket analiz sistemleri, giyilebilir teknolojiler, ivmeölçerler, jiroskoplar ve akıllı telefon tabanlı yazılımlar da yürüyüşün incelenmesinde kullanılmaktadır.

Şekil 2.8’de gösterilen kayıt cihazında ayakkabı üzerinde ve ayak bileğine yakın bir bölgeye yerleştirilen sensörler bulunmaktadır. Bu sensörler aracılığı ile adım frekansı, adım uzunluğu, adım sayısı, yürüyüş döngüsü değişkenliği, yürüyüş hızı, yürüyüş evreleri ile ilgili sayısal veriler, koşu aşamaları, ayak açısı, adım esnasında ayağın yerden yüksekliği, dairesel hareket gibi değerli veriler elde edilebilmektedir. Bu markanın farklı modellerinde sensör sayıları arttırılmış, diz kapağı ve bel bölgesine de sensörler yerleştirilerek daha fazla ve daha detaylı veriler de elde edilebilme imkanı sunulmuştur. Bu veriler: ayak bileği açısı, diz eklemi açısı, kalça eklemi açısı ve bu açılarla ilgili diğer sayısal verilerdir.

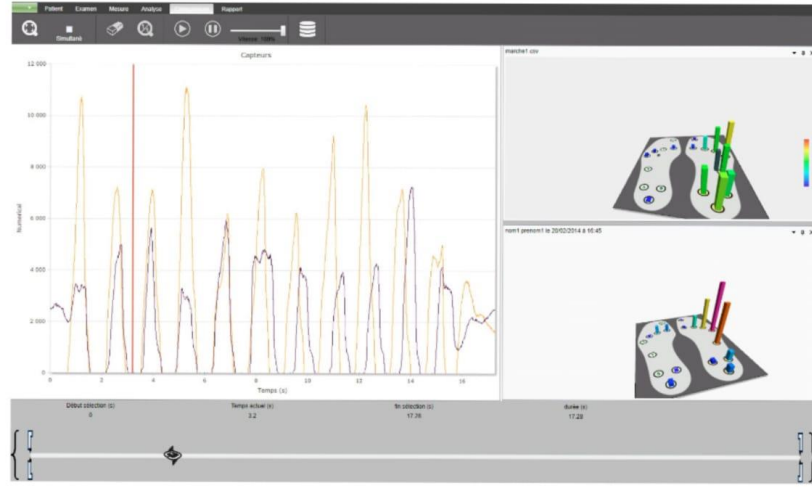
Deneklerin yürüyüş verilerini elde etmek için kullanılan başka bir sistem de kameralı sistemdir (Şekil 2.9). Bu sistem Şekil 2.8’de gösterilmiş olan cihaza göre daha karmaşıktır ancak daha detaylı çıktılar vermektedir. 3 boyutlu veriler elde edilmesine imkân sağlar. Ancak taşınabilir değildir, sistemin kurulumu için geniş ve sabit bir alana ihtiyaç duyulur ve yer değişikliği durumunda kalibrasyon gerektirmesi açısından

dezavantajlıdır. Ancak çıktı olarak hem sayısal hem görsel hem de grafiksel veriler elde edilmesi açısından avantajlıdır.



Şekil 2.9. Yürüyüş analizi sistemlerinde 3 boyutlu kayıt yapan kamera tabanlı sistemlere örnek (mar-systems, 2021).

Parkinson hastalığında incelenen parametrelerden biri de ayak tabanının farklı noktalarından elde edilen kuvvet verileridir. Bu kuvvet verilerinin toplanmasında kullanılan sistem görseli Şekil 2.10'da verilmiştir. Farklı sayıda sensör içeren tabanlı ayakkabı içerisine yerleştirilerek yürüyüş döngüsünün kaydedilmesinde kullanılır. Deneğin yürümesi esnasında sensörlerin bulunduğu her noktadan kuvvet verileri anlık olarak kaydedilir. Böylelikle yürüyüş evrelerinde taban kuvvetlerindeki değişiklikler gözlemlenebilir ve yorumlanabilir.



Şekil 2.10. Ayak tabanı kuvvet verilerinin elde edilmesinde kullanılan sensörlü tabanlık sistemi çıktıları, kullanılışı ve görseli (Medicapture, 2021).

2.4. Parkinson Hastalığının Tespitine Yönelik Yapılan Genel Çalışma Örnekleri

Parkinson hastalığının her ne kadar bilinen tam bir tedavisi olmasa da en azından hastanın yaşam kalitesinin artırılması ve hastalığın ilerlemesinin yavaşlatılması için erken teşhis şarttır. Hastalığın ilerleyen evrelerinde ilaç tedavisi yöntemi uygulansa da belirtilerin baskılanması mümkün olmayabilmektedir. Bu durumda hasta artmış titreme ve katılık, hareketlerde zorlanma vb. zorlukların sonunda artık kendi kendine hareket

edemez, yalnız başına hayatını idame ettiremez hale gelebilmektedir. Belirtileri bu şekilde ağırlaşan ve ilaç tedavisini yanıt alınamayan hastalara son çare olarak beyin pili olarak da bilinen derin beyin stimülasyonu yöntemi uygulanmaktadır. Ancak beyin pili operasyonları hastaların yaş ortalaması da düşünüldüğünde barındırdığı riskler açısından tehlikeli olabilmektedir. Her operasyonda olan enfeksiyon riskine ek olarak, beyin kanaması, görme kaybı gibi kalıcı kusur kalabilme ihtimali vardır. Bu sebeple hastalığın ilerlemeden mümkün olan en erken evrede tespit edilmesi önem arz etmektedir.

Parkinson hastalığının teşhisinde hekimlerin kullandığı UPDRS ölçeklendirmesinden Başlık 2.1’de bahsedilmiştir. UPDRS ölçeği nörologların günümüzde hastalığın teşhisi ve evresinin belirlenmesinde kullandıkları özel bir anket ve puanlandırma sistemidir. Hastanın semptomları hasta ve yakınının beyanına göre puanlandırma anketi şeklinde kaydedilir. Nöroloğun gözlemi de anketin devamında bulunan bölümlerdeki sorulara göre puanlandırma şeklinde kaydedilir. En son tüm puanlar hesaplanarak hastalığın var olup olmadığı ve var ise derecesinin tespiti yapılır. Bu işlemler tamamen manuel olarak ve matbu bir form üzerinde gerçekleştirilir. Bu durum testin zaman almasına neden olmaktadır. Aynı zamanda hesaplama içermesi bakımından da insan hatalarına açık bir durumdur. Ayrıca testin cevaplarının bir kısmı hasta ve hasta yakınının beyanı üzerine kuruludur ve hasta ve/veya yakınının doğru beyan vermeme olasılığı vardır. Bu da anketin ilgili kısmındaki puanlandırmanın yanıltıcı olmasına neden olabilir.

Yukarıda bahsedilen sebepler nedeniyle hastalığın teşhisinde kullanılan UPDRS ölçeğine ek olarak bilgisayar tabanlı sistemlerin kullanılması; daha net, objektif ve nicel veriler ile teşhis yoluna gidilmesi daha faydalı olacağı düşünülmektedir. Bu sebeple Parkinson hastalığı ile ilgilenen araştırmacılar yıllardır hastalığın teşhisi için kullanılabilecek farklı cihazlar ve yöntemler aramaktadırlar. Bu yöntemlere; hastaların ses kayıtlarının alınması ve incelenmesi, hastaların el yazısı kayıtlarının toplanması ve incelenmesi, hastaların elektroensefalografi (EEG) sinyallerinin incelenmesi, hastaların MRG (Manyetik Rezonans Görüntüleme) görüntülerinin incelenmesi, hastaların parmak vuruş hareketlerinin kaydedilmesi ve incelenmesi ve yürüyüş ile ilgili kayıtların alınması ve incelenmesi örnek verilebilir.

2018 yılında İtalyan bilim insanları 37 Parkinson hastası 38 sağlıklı birey arasında el yazısı veri kayıtları ile makine öğrenmesi yöntemleri kullanarak hasta-sağlıklı ayrımı yapabilen bir algoritma üzerine çalışmışlardır. Çalışmada makine öğrenmesi sınıflandırıcılarından KNN (k-nearest neighbors), SVM (support vector machines), NB (Naive Bayes), RF (Random forest), ADA (AdaBoost) algoritmaları kullanılmıştır. Sonuçta en yüksek %73,38 doğruluk oranına RF sınıflandırıcı ile ulaşılmıştır (Impedovo, Pirlo & Vessio, 2018).

50 sağlıklı birey ve 82 Parkinson hastası ile yapılan farklı bir çalışmada ise deneklerin işaret ve başparmaklarına yerleştirilen sensörler ile veri toplanmıştır. Deneklerden başparmaklarını ve işaret parmaklarını kendi belirledikleri hızda birbirine dokundurmaları istenmiştir. Ekip, kullandıkları sınıflandırma algoritması ile hastalık varlığını %93 oranında doğru tahmin edebilmiştir (Alty, Cosgrove, Lones, Smith, Possin, Schuff & Jamieson, 2016).

Literatürde ses kayıtlarından Parkinson hastalığı ayrımı yapmayı amaçlayan çalışmalar da bulunmaktadır. Örnek bir çalışmada 23 Parkinson hastası 8 sağlıklı bireyin ses kayıtları alınmıştır. Hastalara kayıtlar esnasında aynı konuşma testleri yapılmıştır. Sonuç olarak SVM sınıflandırıcı algoritması ile %91,4'lük bir başarı elde edilmiştir (Little, McSharry, Hunter, Spielman, & Ramig, 2008).

27 Parkinson hastası ve 49 sağlıklı birey ile gerçekleştirilen bir başka çalışmada, hasta sağlıklı ayrımını %78 doğru gerçekleştirebilen bir algoritma üretilmiştir (Adams, 2018).

Parkinson hastalığının tespitinde kullanılan bir başka veri türü de postural denge verileridir. 67 (31 erkek, 36 kadın) katılımcı ile yapılmış bir araştırmada Parkinson hastası olan ve olmayan sağlıklı bireylerden kendilerini rahat hissettikleri bir pozisyonda 90 saniye süresince özel bir kuvvet ölçüm platformu üzerinde hareketsiz durmaları istenmiş ve bu esnada kayıt alınmıştır. Platform; hastanın vücut ağırlık merkezinin konum-zaman bilgilerini anteroposterior (merkezden öne-arkaya doğru) ve mediolateral (merkezden sağa ve sola doğru) olarak kaydetmektedir. Sınıflandırma algoritması olarak LR (Lojistik regresyon) algoritması kullanılmıştır. Sonuç olarak, doğruluk oranı %83 olan bir hasta

sağlıklı sınıfı ayırımı yapılmıştır (Chou, Chou, Chen & Yang, 2020). 55 Parkinson hastası, 55 sağlıklı birey ile yapılan bir başka denge çalışmasında ise; deneklerden kuvvet ölçüm platformuna çıkmaları ve 30 saniye boyunca gözleri kapalı, 30 saniye boyunca da gözleri açık ve sabit bir noktaya bakarak rahat bir pozisyonda ayakta durmaları istenmiştir. Bu esnada kuvvet ölçüm platformu hastaların vücut ağırlık merkezlerindeki konum değişikliklerini milimetre cinsinden kaydetmiştir. Yapılan istatistiksel analizde hasta ve sağlıklı bireyler arasında gözün açık ve gözün kapalı olması durumlarında vücut ağırlık merkezi konum değişikliklerinin anlamlı derece farklı olduğu tespit edilmiştir (Błaszczyk, Orawiec, Duda-Kłodowska & Opala, 2007).

Çizelge 2.1. Parkinson hastalığının tespitinde farklı verilerle yapılmış çalışma örnekleri özeti.

Çalışma	Deney Grubu	Mod	Yöntem	Sonuçlar
Impedovo vd. (2018)	37 Parkinson Hastası, 38 sağlıklı	El yazısı	Makine öğrenmesi	RF sınıflandırıcı ile %73.38 doğruluk
Alty vd. (2016)	49 Parkinson Hastası, 41 sağlıklı	Parmak vuruş hareketi	Optimizasyon Algoritması	%93 doğruluk
Little vd. (2008)	23 Parkinson hastası, 8 sağlıklı	Ses kaydı	Makine öğrenmesi	SVM sınıflandırıcı ile %91,4 doğruluk
Adams (2018)	27 Parkinson hastası, 49 sağlıklı	Tuş vuruşu	Makine öğrenmesi	%78 doğruluk
Chou vd. (2020)	67 katılımcı	Postural denge	Makine öğrenmesi	LR %83
Błaszczyk vd. (2007)	55 Parkinson hastası, 55 sağlıklı	Postural Denge	İstatistiksel analiz	İstatistiksel ayırma anlamlı sonuçlar elde edilmiştir.
Combs, Diehl, Filip & Long (2014)	88 Parkinson hastası	Yürüyüş	Anket ve istatistiksel analiz	İleri evrede hastaların yavaş yürüdüğü ispatlanmıştır.
Kimmeskamp & Hennig (2001)	24 Parkinson hastası, 24 sağlıklı	Taban kuvveti	İstatistiksel analiz	Taban kuvvet noktalarında sağlıklı bireylerden farklı olarak

				topukta kuvvette azalma, ön ayak bölgesinde artma gözlenmiştir.
Arora vd. (2018)	334 Parkinson hastası, 104 iRBD hastası, 84 sağlıklı	Ses, yürüyüş, denge, reaksiyon zamanı, titreme	Makine öğrenmesi	Parkinson-sağlıklı ayrımı yapabilen sistemde %83,3 kesinlik oranına ulaşılmıştır.
Abdulhay, Arunkumar, Narasimhan, Vellaiappan & Venkatraman (2018)	93 Parkinson hastası, 73 sağlıklı	Taban kuvveti	Makine öğrenmesi	%94,8 doğruluk
Balaji, Brindha & Balakrishnan (2020)	93 Parkinson hastası, 73 sağlıklı	Yürüyüş	Makine öğrenmesi	Hem hastalık tespiti hem de evre tayini yapılmıştır. %76-100 aralığında doğruluk
Manap, Tahir, Yassin & Abdullah (2011a)	12 Parkinson hastası, 20 sağlıklı	Yürüyüş	Makine öğrenmesi	%87,50 doğruluk
Manap, Tahir, & Yassin (2011b)			Makine öğrenmesi	%95,63 doğruluk
Tahir & Manap (2012)			Makine öğrenmesi	SVM sınıflandırıcı ile %100 doğruluk
Shetty & Rao (2016)	15 Parkinson, 20 Huntington, 13 ALS hastası, 16 sağlıklı		Makine öğrenmesi	Farklı hastalıkların tespiti. Genel doğruluk oranı %83,33, Parkinson hastalığının tespit oranı: %75
Mc Ardle, Del Din, Galna, Thomas & Rochester (2020)	Alzheimer, DLB, Parkinson hastası, 80 katılımcı	Yürüyüş	İstatistiksel analiz	İstatistiksel ayırmda anlamlı sonuçlar elde edilmiştir.

2.5. Parkinson Hastalığının Yürüyüş ile Tespitine Yönelik Çalışma Örnekleri

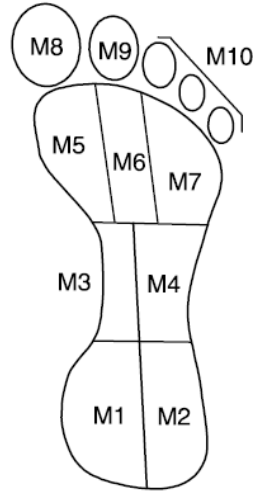
2014 yılında ABD’li bilim insanlarının yaptıkları bir araştırmada, 88 Parkinson hastası bireye farklı yürüyüş testleri ve anketler uygulanmıştır. 88 denek Parkinson hastalığı teşhisi almış kişilerdir ancak hastalık evreleri farklıdır. Yürüyüş testleri: 10 metre hastaların kendilerinin belirledikleri hızda yürüme (comfortable 10-m walk tests (CWT)), 10 metre yürüyebildikleri maksimum hızda yürüme (fast 10-m walk tests (FWT)), 6 dakikalık mini yürüme testi (6-min walk test (6MWT)) ve mini denge (mini balance evaluations systems test (mini-BEST Test)), düşme korkusu (fear of falling), aktiviteye özgü denge güven ölçeği (the Activity-Specific Balance Confidence Scale (ABC)) şeklindedir. Anketler hastalar arasında yürüme esnasında düşme korkusu yaşayan/yaşamayan ve hastaların günlük aktivitelerindeki denge öz yeterliliklerini saptamak için hazırlanmış 16 sorudan oluşmaktadır. Sonuçta CWT, FWT ve 6MWT testlerinde hastalık derecesi ileri seviyede olan hastalar, düşük seviyede olan hastalara göre anlamlı derecede yavaş yürümüşlerdir (Combs vd., 2014). Bu sonuç bizlere Parkinson hastalığının yürüyüş hızı üzerinde olumsuz bir etkisinin olduğunu bir kez daha kanıtlamıştır.

Parkinson hastalığının tespitinde kullanılan diğer belirtilerden olan ayak tabanındaki basınç noktalarındaki değişiklikler, denge problemleri, düşme korkusu ve ayak sürüme de araştırmacılar tarafından incelenmiştir. Bu araştırmalardan biri 2001 yılında Kimmeskamp & Hennig tarafından yapılmıştır. Araştırmacılar 24 Parkinson hastası ve 24 sağlıklı bireyin yürüyüş esnasında ayak tabanı basınç bölgelerindeki yükleri kaydederek incelemişlerdir. Bu kayıtlar için Pedar marka sensörlü tabanlık kullanılmıştır (Şekil 2.11). Tabanlık 99 farklı noktadan basınç ölçümü yapmaktadır.



Şekil 2.11. Ayakkabı tabanlığına yerleştirilen 50 Hz frekanslı, 99 farklı sensör bulunan tabanlık-Pedar marka (Novel, 2021).

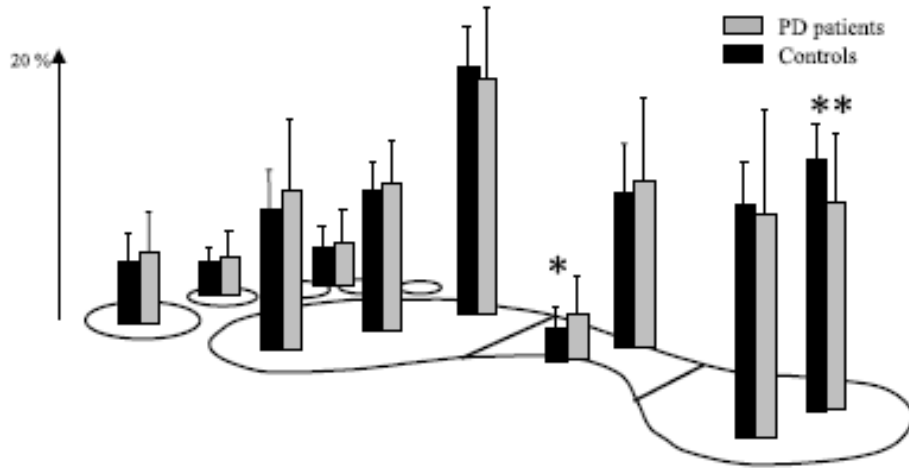
- M1: medial heel
- M2: lateral heel
- M3: medial midfoot
- M4: lateral midfoot
- M5: metatarsal head I
- M6: metatarsal head II
- M7: metatarsal heads III-V
- M8: hallux
- M9: second toe
- M10: lateral toes



Şekil 2.12. Ayak tabanı basınç bölgelerinin 10 farklı anatomik bölgeye ayrılması (Kimmeskamp & Hennig, 2001).

Çalışmada ayak tabanı 10 farklı anatomik alana bölünmüştür ve her alan için ayrı ortalama basınç değerleri hesaplanmış ve kullanılmıştır (Şekil 2.12). M1, M2: Arka ayak (topuk bölgesi); M3,M4: Orta ayak bölgesi; M5, M6, M7, M8, M9 ve M10: ön ayak bölgesidir.

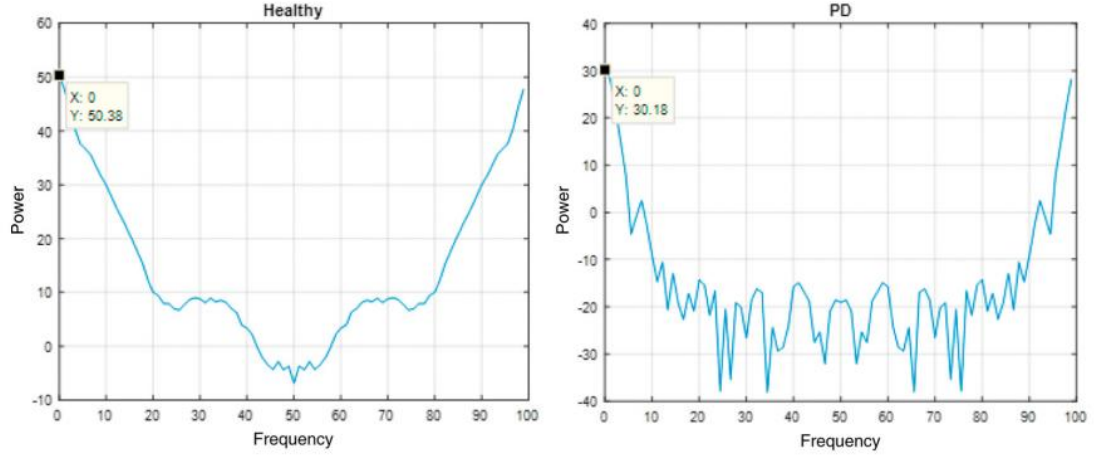
Çalışmaya katılan hasta ve sağlıklı bireylerden kendilerini rahat hissettikleri bir hızda 11 metre yürümeleri istenmiştir. Elde edilen sonuçlara göre Parkinson hastası grupta, topuk bölgesindeki basınç değerlerinin, sağlıklı bireylere göre anlamlı derecede düşük olduğu tespit edilmiştir. Hasta grubunda topuk bölgesinden, orta ve ön ayak bölgesine doğru bir yük kayması olduğu sonucuna varılmıştır. Sağlıklı bireyler yürüyüş döngülerini başlatırken adım atış esnasında ilk olarak topuğun yere temas ettirirler ve bunu orta ve ön ayak bölgesi (ayak parmakları) takip eder. Ancak Parkinson hastaları sağlıklı bireylerin aksine adım atma sırasında ilk olarak ön ve/veya orta ayak bölgesini yere temas ettirmişlerdir. Özellikle ileri evrelerdeki hastalarda hastalar ayak bileğini hareket ettirmeden ve ayak tabanını yere vurarak tüm ayak tabanı bölgelerini neredeyse aynı anda yere temas ettirerek adım atmışlardır. Hasta bireyler ile sağlıklı bireyler arasındaki bir diğer farklılık ise; tam basma evresinde hastaların sağlıklı bireylere göre taban basınçlarının parmak uçlarına doğru; ön ve orta bölgelere kaymasıdır. Yani hastalığın etkilerinden biri de topuk bölgesinde yük azalması, bu yükün parmaklara doğru kaymasıdır denilebilir (Şekil 2.13). Bu sonuç Parkinson hastalarının tipik olarak öne doğru kambur durmalarını ve yürüyüş esnasında ayaklarını yere vurarak ve öne eğik pozisyonda yürümelerini de açıklamaktadır (Kimmeskamp & Hennig, 2001).



Şekil 2.13. Hasta ve sağlıklı bireylerin ayak basınç noktaları ortalamalarının karşılaştırılması (Kimmeskamp & Hennig, 2001).

Oxford Üniversitesi'nden bilim insanlarının 2018 yılında yaptıkları bir çalışmada 334 Parkinson hastası, 104 iRBD (idiopathic REM sleep behavior disorder (idiyopatik REM uyku davranış bozukluğu)) hastası ve 84 sağlıklı bireyden, ses, yürüyüş denge, reaksiyon zamanı, istirahat tremoru (dinlenme halindeki titreme) ve postural tremor (hareket halindeyken titreme) verileri kaydedilmiştir. Çalışma Parkinson hastalığında çoklu verilerin kullanılması açısından özgündür. Yürüyüş ve denge ile ilgili kısımda; kayıt için deneklerin ceplerinde bulunan akıllı telefonun atalet (eylemsizlik) ölçüm birimleri (IMU'lar) ve üç eksenli ivmeölçer sensöründen faydalanılmıştır. Aynı çalışmada yürüyüş verileri de benzer şekilde hasta ve kontrol gruplarının ceplerindeki akıllı telefon ile 20 metre ileri ve geri yürümeleri istenilerek toplanmıştır. Sonuçta Parkinson hastaları ile sağlıklı bireyleri birbirinden ayıran sistemde postural titreme en belirleyici veriler olarak belirtilmiştir. Parkinson hastası-sağlıklı ayrımı yapan sistemde denge ve yürüyüş ile ilgili veriler ayırma anlamlı katkılar sağlamamıştır. Ancak kadınlar arasında yapılan ayırma; iRBD-sağlıklı ve iRBD-Parkinson hastası ayırımında hem yürüyüş hem de denge verileri anlamlı katkı sağlamıştır (Arora vd., 2018).

Yürüyüş ve ayak tabanı kuvvetleri ile Parkinson hastalığının tespiti ile ilgili yapılan bir başka çalışmada da yine hasta ve sağlıklı bireylerin ayak tabanına kuvvet sensörleri yerleştirilerek veriler elde edilmiş ve bu verilere makine öğrenmesi yöntemleri uygulanarak sınıflandırma yapılmıştır. Çalışmada 93 Parkinson hastası ve 73 sağlıklı birey katılımcı olmuştur. Veriler toplam 8 farklı sensör içeren bir çift tabanlı olarak kullanılarak toplanmıştır. Sayısal veriler incelendiğinde Parkinson hastası bireylerin taban güç frekanslarının sağlıklı bireyler ile son derece farklı olduğu saptanmıştır (Şekil 2.14). Sonuçta Parkinson hastaları ile sağlıklı bireyleri %94,8 oranında doğru sınıflayabilen bir sınıflandırma algoritması geliştirilmiştir. Kullanılan sınıflandırma yöntemi SVM'dir. Ayırma hem topuk basınç noktaları değişiklikleri hem de adım sayısı, duruş zamanı, salınım zamanı gibi parametrelerin etkili olduğu gözlemlenmiştir (Abdulhay vd., 2018).



Şekil 2.14. Sağlıklı ve Parkinson hastası bireylerin kuvvet sinyalinin güç spektral yoğunluğu karşılaştırması (Abdulhay vd., 2018).

Yapılan farklı bir başka çalışmada ise makine öğrenmesi yöntemleri kullanılarak Parkinson hastalığının hem tespiti hem de hastalığın evre sınıflandırması yapılmıştır. Araştırmada 93 Parkinson hastası ve 73 sağlıklı bireyden temel ve kinetik yürüyüş verileri kaydedilmiştir. Kullanılan temel yürüyüş verileri: adım süresi, duruş süresi, salınım süresi, salınım/duruş oranı, çift adım süresi, kadans, hız, adım uzunluğu, çift adım uzunluğudur. Kinetik verilerde ise Şekil 2.11.'de bulunan cihaza benzer bir cihaz ile elde edilen ve Şekil 2.12.'te yer alan ancak farklı olarak 8 anatomik bölgeden oluşan taban kuvveti değerleri kullanılmıştır. Çalışmada seçilen hasta bireylerin 10'u ileri evrede hasta iken, 28'i orta evre hasta ve 55'i de başlangıç derecesindedir. Verilere SVM, karar ağaçları, topluluk sınıflandırma ve Naive Bayes sınıflandırma algoritmaları uygulanmıştır. Sonuçta %76 ile %100 aralığında doğru, %0 ile %100 aralığında hassas sonuçlar elde edilmiştir (Balaji vd., 2020).

Bu çalışmada kullanılan verileri toplayan ve kullanan Malezya MARA Teknoloji Üniversitesi'nden Prof. Nooritawati Md Tahir ve ekibi, elde ettikleri veriler ile 2011 ve 2012 yıllarında toplamda 3 farklı çalışma yapmışlardır. Veriler 12 Parkinson hastası ve 20 sağlıklı bireyden alınan yürüyüş verileridir. Temel yürüyüş verileri ve kinematik veriler 37 yansıtıcı sensör ve 6 kızılötesi kamera kullanılarak elde edilmiştir. Yüzey tepki kuvvetleri olarak da isimlendirilen kinematik veriler ise 200 Hz örnekleme hızına sahip iki kuvvet plakası ile elde edilmiştir.

Malezyalı ekibin yaptığı ilk çalışma Haziran 2011 yılında yayınlanmıştır. Bu çalışmada temel, kinetik ve kinematik öznitelikler olmak üzere 3 ana başlıklı toplamda 21 öznitelik kullanılmıştır. 21 öznitelik Çizelge 2.2’de ayrıntılı olarak verilmiştir. Veriler ANN (Artificial Neural Network (Yapay sinir ağları (YSA)) algoritması kullanılarak sınıflandırılmıştır. Çapraz doğrulama yapılırken deneklere ait veriler 4 gruba bölünerek kullanılmıştır. 1. Grupta 6 sağlıklı, 2 Parkinson hastası; 2. Grupta 5 sağlıklı, 3 Parkinson hastası; 3. Grupta 5 sağlıklı, 3 Parkinson hastası, 4. Grupta 4 sağlıklı, 4 Parkinson hastası bulunmaktadır. Gruplardan biri test seti olarak kullanılırken, üçü eğitim seti olarak kullanılmıştır. Her bir grup test seti olacak şekilde 4 çapraz doğrulama yapılmıştır. Her bir grubun test seti olduğu sınıflandırma sonuçları ayrı ayrı kaydedilmiş ve ortalamaları alınarak nihai sonuç kaydedilmiştir. Ayrıca temel yürüyüş verileri, kinetik yürüyüş verileri ve kinematik yürüyüş verileri önce tek tek sonra ikili olarak en son da üçü birlikte olarak farklı varyasyonlarla sınıflandırılmıştır. Buradaki amaç sonuçları karşılaştırıp yorumlayarak temel, kinematik ve kinetik verilerden hangisi ya da hangilerinin sınıflandırıcı başarısını arttırdığını tespit etmektir. Sınıflandırma sonuçları incelendiğinde en yüksek ortalama doğruluk değeri %87,50 olarak kaydedilmiştir. Temel yürüyüş verileri, kinetik yürüyüş verileri ve kinematik yürüyüş verilerinin tek tek, ikili olarak ve üçü birlikte kullanıldığı farklı sınıflandırma varyasyonlarının sonuçları karşılaştırıldığında yalnızca kinematik verilerin yalnızca kinetik ve yalnızca temel yürüyüş verilerinin kullanıldığı varyasyonlara göre daha yüksek doğruluk oranı ile sınıflandırma yaptığı görülmektedir. Bu durum yürüyüşteki eklem açılarının Parkinson hastası bireylerde daha küçük adımlarla yürümekten kaynaklı olarak azalmış olması ve bu nedenle kinematik verilerin daha belirleyici olabileceği yönünde yorumlanmıştır. Ek olarak en yüksek sınıflandırma doğruluk oranı olan %87,50 üç veri türünün tamamının kullanıldığı sınıflandırıcıda elde edilmiştir (Manap d. 2011a).

Çizelge 2.2. Bu çalışmanın da temelini oluşturan Prof. Nooritawati Md Tahir ve ekibinin yürüyüş verileri hakkında detaylar

Temel yürüyüş verileri	Kinematik yürüyüş verileri	Kinetik yürüyüş verileri
Adım zamanı	Topuk vuruşunda kalça açısı	F_{y1} (frenleme aşamasında maximum yatay topuk kuvveti)
Kadans	Ayakucunda kalça açısı	F_{y2} (maximum yatay itme kuvveti)
Adım uzunluğu	Maximum fleksiyonda kalça açısı	F_{z1} (maximum dikey topuk kuvveti)
Yürüme hızı	Minimum extensionda kalça açısı	F_{z2} (minimum dikey orta duruş kuvveti)
	Topuk vuruşunda diz açısı	F_{z3} (maximum dikey itme kuvveti)
	Ayakucunda diz açısı	
	Maximum fleksiyonda diz açısı	
	Minimum extensionda diz açısı	
	Topuk vuruşunda bilek açısı	
	Ayakucunda bilek açısı	
	Maximum fleksiyonda bilek açısı	
	Minimum extensionda bilek açısı	

Malezyalı ekibin yaptığı ikinci çalışma Aralık 2011’de yayınlanmıştır. Bu çalışma da Çizelge 2.2’de ayrıntıları verilmiş olan veri seti kullanılmıştır. Haziran 2011’deki çalışmalarından farklı olarak burada GRF normalizasyon tekniği kullanılmıştır (Denklem 2.2). Burada GRF (ground reaction forces) olarak nitelendirilen veriler esasen kinetik yürüyüş verileridir. Hastalar ve sağlıklılar arasındaki patolojik yürüyüş farklılıklarının belirlenmesi için verilere bağımsız t-test ve Pearson korelasyon testi olmak üzere iki istatistiksel analiz uygulanmıştır. Bu iki istatistiksel analiz yönteminden 4 öznitelik belirleyici olarak ayırt edilmiştir. Seçilen bu dört öznitelik: adım uzunluğu, yürüme hızı, diz açısı ve dikey kuvvet öznitelikleridir. Yapılan bu analizlere ek olarak bu çalışmada

Kfold çapraz doğrulama yöntemi ve yine ANN sınıflandırma algoritması kullanılmıştır. ANN sınıflandırıcısında temel, kinetik ve kinematik yürüyüş verilerinin tamamının kullanıldığı sınıflandırıcı doğruluk oranı 87,25 olarak elde edilmiştir. Bu çalışmada istatistiksel analiz yöntemleri ile elde edilen 4 öznitelik kullanıldığında doğruluk oranı %95,63'e yükselmiştir (Manap vd., 2011b).

$$GRF(\%) = \frac{GRF(N)}{Vücut \text{ Ağırlığı}} \times 100 \quad \text{Denklem 2.2.}$$

$$X_{intra} = \frac{X_i}{X_{\max(\text{grup içi})}} \quad \text{Denklem 2.3.}$$

$$X_{inter} = \frac{X_i}{X_{\max(\text{tüm veri})}} \quad \text{Denklem 2.4.}$$

Malezyalı ekibin yaptığı üçüncü çalışma ise 2012 yılındadır. Aynı veri seti üzerinde bu kez farklı normalizasyon yöntemleri ve farklı bir sınıflandırma yöntemi uygulamışlardır. GRF normalizasyon (Denklem 2.2.), intra grup normalizasyon (Denklem 2.5), inter grup normalizasyon (Denklem 2.6) olmak üzere 3 farklı normalizasyon yöntemi uygulanmıştır. İlk çalışmalarında olduğu gibi veri seti dörde bölünmüş ve çapraz geçişleme uygulanmıştır. İnter grup normalizasyon ve intra grup normalizasyon verileri ayrı ayrı sınıflandırıcılarda çalıştırılmıştır. Sınıflandırıcı algoritması olarak SVM ve yine ANN kullanılmıştır. İnter grup normalizasyonda temel yürüyüş verileri %100'lük bir sınıflandırma başarısı göstermiştir. Temel, kinetik ve kinematik yürüyüş verilerinin tamamı kullanılarak yapılan sınıflandırmada en yüksek doğruluk değerine intra grup normalizasyon ile SVM sınıflandırıcıda ulaşılmıştır (Tahir & Manap, 2012).

Parkinson hastalığının tespitinde yürüyüş verilerinin kullanıldığı çalışmalar yukarıda özetlenmiştir. Ancak yürüyüş verileri kullanılarak yalnızca hasta-sağlıklı ayrımı ya da hastalık evresi ayrımı yapmanın dışında da çalışmalar bulunmaktadır. Arora vd. nin yaptıkları çalışmada olduğu gibi farklı hastalıkların ayrımında da benzer yöntemler kullanılabilir. Bu durum özellikle belirtileri benzer olan, ayırt edilmekte zorlanılan

hastalıkların tespitinde de yürüyüş analiz yöntemlerinin ve makine öğrenmesi gibi yöntemlerin kullanılabilceğini göstermektedir.

2016 ağustos ayında Hindistan'da yapılan bir çalışmada 15 Parkinson hastası, 20 Huntington hastası, 13 ALS hastası ve 16 sağlıklı bireyden yürüyüş verileri toplanmıştır. Çalışmanın amacı; Parkinson hastalığını farklı iki nörolojik hastalık olan ALS ve Huntington hastalığından ayırt edebilmektir. Yürüyüşlerden elde edilen verilerde 7 farklı öznitelik bulunmaktadır ve bu öznitelikler kullanılarak makine öğrenmesi yöntemi ile sınıflandırma yapılmıştır. Sonuçta; SVM sınıflandırma yöntemi ile %83,33'lük genel bir doğruluk oranı elde edilmiş ve Parkinson hastalığını %75 oranında doğru tespit edebilen bir sistem geliştirilmiştir (Shetty & Rao, 2016).

Benzer bir çalışma da 2020 yılında İngiltere'de yapılmıştır. Çalışmada Alzheimer Hastalığı, Lewy cisimcikli demans (Dementia with Lewy bodies (DLB)) hastalığı ve Parkinson hastalığı teşhisi almış toplam 80 kişiden giyilebilir sensörler aracılığıyla yürüyüş verileri elde edilmiştir. Yürüyüşün hız, değişkenlik, ritim, asimetri ve postural kontrol özniteliklerinden faydalanılarak istatistiksel bir analiz yapılmıştır. Yapılan analiz sonucunda farklı hastalıklara sahip bireylerin yürüyüş verilerinin anlamlı derecede farklı olduğu tespit edilmiştir (Mc Ardle vd., 2020).

BÖLÜM 3

MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE YÜRÜYÜŞ VERİLERİ KULLANILARAK PARKİNSON HASTALIĞININ TESPİTİ

3.1. Makine Öğrenmesi Hakkında Genel Bilgiler

“Öğrenme” sözcüğünün Türk Dil Kurumu’ndaki tanımı: “Bireyin yaşantıları ve çevresi ile etkileşimi yoluyla bilişsel, duyuşsal ve devimsel alanda kısmen kalıcı, gözlenip ölçülebilir davranış kazanması ya da eski davranışlarını değiştirme süreci.” şeklindedir. Öyleyse makine öğrenmesi kavramı da “belirli bir veri kümesi vasıtası ile makinenin çıkarımlar, tanımlar, tahminler yapabilmesi” şeklinde tanımlanabilir.

Her gün daha da ilerleyen teknoloji, sanayileşme ve bilgisayar kullanımı veri sayısında önemli bir artışın yaşanmasına sebep olmuştur. Bu durumun paralelinde verileri toplama ve verilere ulaşabilme yöntemleri de gelişmiştir. Ancak bu büyük veri yığınlarının anlamlı hale getirilmesi ve verimli kullanılabilmesi ihtiyacını da doğurmuştur. Bu ihtiyaç doğrultusunda veri madenciliği, makine öğrenmesi, yapay zekâ ve derin öğrenme gibi yeni yöntemler geliştirilmektedir.

Makine öğrenmesi algoritmaları bilgisayara geçmiş veriler kullanılarak tecrübe kazandırma veya bir başka deyişle öğrenme yolu ile sonuç analizinde bulunabilme,

gelecek hakkında tahmin ve/veya modelleme yapabilme görevlerini gerçekleştirilebilmeye imkân tanımaktadır. Makine öğrenmesiyle bilgisayara daha önceki örneklerden elde edilmiş deneyimlerin öğretilmesi sağlanır. Büyük hacimde verilerde makine öğrenmesi teknikleri uygulanarak değerli sonuçlar elde edilebilir. Günümüzde özellikle karar destek sistemlerine ihtiyaç duyulan alanlarda makine öğrenmesi teknikleri sıkça kullanılmaktadır. Bu konuda önerilen çok sayıda algoritma ve yaklaşım bulunmaktadır. Bu yaklaşımların kimi tahmin (prediction) ve kestirim (estimation) kimi de sınıflandırma (classification) yapabilme üzerinedir (Orakcı, Ciylan, Kök & Sevri, 2019).

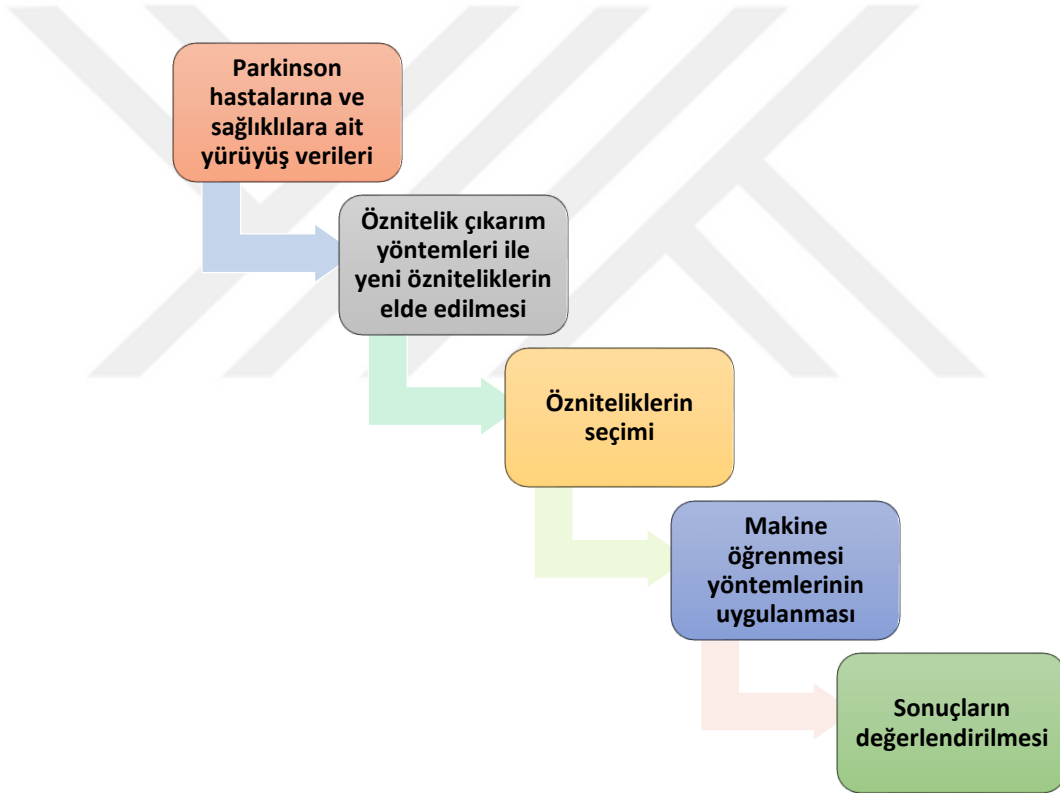
Veri madenciliği ve makine öğrenmesi kavramları arasında doğrudan bir ilişki bulunmaktadır. Makine öğrenmesi yöntemlerinin büyük verilere uygulanması veri madenciliği olarak ifade edilebilir. Yani makine öğrenmesi; veri madenciliğinin uygulama sürecidir denilebilir. Seçilen veri seti üzerine seçilen makine öğrenmesi tekniği uygulanır ve bazı sonuçlar elde edilir. İşte burada makine öğrenmesi; öğrenme metotları, tahminler, tanımlar oluşturmak için kullanılır ve bu metot, tahmin ve tanımların maksimum performans ile nasıl çıkarılabileceği ile ilgilenir. Ancak veri madenciliği makine öğrenmesi ile ortaya çıkan bilgi ve bu bilgilerin değerlendirilmesi ile uğraşmaktadır. Yani makine öğrenmesi bir uygulama süreci iken, veri madenciliği bu uygulamanın sonuçları ile ilgilenir. Öyleyse veri madenciliği çıktılarının başarısı, makine öğrenmesi algoritmalarının başarısına doğrudan bağlıdır şeklinde bir yorum yapılabilir (Orakcı vd., 2019).

Makine öğrenmesi kullanılacak verilerin toplanma şekli ve format değişiklikleri ile ilgilenmez. Yalnızca verilerden elde edeceği tahminleri ya da sınıfları en iyi performansla nasıl çıkaracağı ile ilgilenir. Ayrıca veri setine eklenen her yeni veri ya da verilerde yapılan tüm değişiklikler performansın daha da gelişmesini sağlar yani öğrenme performansını artırır denilebilir (Dalyan, 2006).

Makine öğrenmesi algoritmalarının uygulanabilirliği için öncelikle verilere ön inceleme yapılması şarttır. Verilerin kullanılabilmesi için uygun formatta olması, gereksiz olanların veri setinden çıkarılması, eksik olan verilerin tamamlanması ve verilere ölçeklendirme yapılması gereklidir.

Makine öğrenmesi algoritmalarında ilgili tahmin ya da modelleme yapabilmek için; her örneğe ait kullanılabilir her türlü özel bilgiye *öznitelik (feature)* denir. Çıkarım için özneliğin yanında, verilerin sınıf bilgileri de gereklidir. Makine öğreniminde kullanılan modele öğretebilmek için gerekli olan verilere *eğitim verileri*, karşılaştırma için ayrılan verilere ise *test verisi* adı verilir (Kurt, 2019). Burada eğitim verileri öğrenme olayının gerçekleştirilmesi için gerekli iken, test verileri eğitilen sistemin performansının belirlenmesi için gereklidir.

Bu tez çalışmasının makine öğrenim süreci Şekil 3.1’de gösterilmiştir.



Şekil 3.1. Bu tez çalışmasının makine öğrenim süreci.

3.2. Veri Kümelerinin Elde Edilmesi

Bu çalışmada Malezya MARA Teknoloji Üniversitesi’nden Prof. Nooritawati Md Tahir ve ekibinin 12 Parkinson hastası ve 20 sağlıklı bireyden elde ettiği yürüyüş verileri kullanılmıştır. Her kişiye ait yalnızca bir kayıt bulunmaktadır. Kullanılan verilere ait yaş ve cinsiyet ile ilgili bazı bilgiler Çizelge 3.1’de paylaşılmıştır.

Ham veri incelendiğinde hem hasta hem de sağlıklı bireyler için toplamda 21 öznelik bulunmaktadır. Ancak öznelik çıkarım yöntemleri uygulanarak bu sayı bu çalışmada arttırılmıştır.

Çizelge 3.1. Bu çalışmada kullanılan veri kümeleri hakkında bilgiler.

	Parkinson Hastaları	Sağlıklılar
Katılımcı Sayısı	12	20
Kadın Katılımcı Sayısı	4	8
Erkek Katılımcı Sayısı	8	12
Yaş Aralığı	52-75	19-32
Yaşta Standart Sapma	8,26	3,11
Ortalama Yaş	61,42	21,75

3.3. Öznelik Çıkarımı

Ham veriler tüm katılımcılar için 21 öznelik içermektedir. Bu verilere katılımcıların demografik verileri, yaş, cinsiyet, boy ve kilo gibi demografik verileri dâhil değildir. Ham veride yürüyüş verileri temel, kinetik ve kinematik olmak üzere üç bölüme ayrılmıştır.

Kinetik veriler: F_{y1} (frenleme aşamasında maximum yatay topuk kuvveti), F_{y2} (maximum yatay itme kuvveti), F_{z1} (maximum dikey topuk kuvveti), F_{z2} (minimum dikey orta duruş kuvveti), F_{z3} (maximum dikey itme kuvveti) olmak üzere 5 öznelikten oluşmaktadır.

Kinematik veriler topuk vuruşunda kalça açısı, ayakucunda kalça açısı, maximum fleksiyonda kalça açısı, minimum extensionda kalça açısı, topuk vuruşunda diz açısı, ayakucunda diz açısı, maximum fleksiyonda diz açısı, minimum extensionda diz açısı, topuk vuruşunda bilek açısı, ayakucunda bilek açısı, maximum fleksiyonda bilek açısı, minimum extensionda bilek açısı olmak üzere 12 öznelikten oluşmaktadır.

Temel yürüyüş verileri ise adım zamanı, kadans, adım uzunluğu, yürüme hızı olmak üzere 4 öznelikten oluşmaktadır.

Makine öğrenmesi algoritmalarının performansını arttırmak ve daha ayrıntılı sonuçlar elde edebilmek adına 21 olan öznelik sayısı farklı öznelik çıkarım algoritmaları uygulanarak 152'ye çıkarılmıştır. Öznelik çıkarımı ve sınıflandırmada MATLAB (The MathWorks, Inc.) programı kullanılmıştır. Çizelge 3.2'de 152 öznelik tablo halinde verilmiştir.

Çizelge 3.2. Yürüyüş verileri ve çıkarılan öznelikler.

NO	ÖZNELİK	NO	ÖZNELİK
1	kinetic_fy1	77	Katz_kinematic_knee'
2	kinetic_fy2	78	SampEnt3_kinematic_knee'
3	kinetic_fz1	79	multiscaSmplEnt_kinematic_knee'
4	kinetic_fz2	80	HarmMean_kinematic_knee'
5	kinetic_fz3	81	RMS_kinematic_knee'
6	min_kinetic	82	Variance_kinematic_knee'
7	max_kinetic	83	trim_kinematic_knee'
8	ort_kinetic	84	Mink_kinematic_knee'
9	medyan_kinetic	85	Maxk_kinematic_knee'
10	std_kinetic	86	Coefficient_Variation_kinematic_knee'
11	aralik_kinetic	87	approxEnt_kinematic_knee'
12	entropy_kinetic	88	apen_kinematic_knee'
13	carpiklik_kinetic	89	fast_apen_kinematic_knee'
14	basiklik_kinetic	90	covariance_kinematic_knee'
15	kolmogorov_kinetic	91	dfa_kinematic_knee'
16	Higuc_kinetic	92	kinematic_ankle_HeelStrike'
17	Katz_kinetic	93	kinematic_ankle_ToeOff'
18	SampEnt3_kinetic'	94	kinematic_ankle_MaxFlexion'
19	multiscaSmplEnt_kinetic'	95	kinematic_ankle_MaxExtension'
20	HarmMean_kinetic'	96	min_kinematic_ankle'
21	RMS_kinetic'	97	max_kinematic_ankle'
22	Variance_kinetic'	98	ort_kinematic_ankle'
23	trim_kinetic'	99	medyan_kinematic_ankle'
24	Mink_kinetic'	100	std_kinematic_ankle'
25	Maxk_kinetic'	101	aralik_kinematic_ankle'
26	Coefficient_Variation_kinetic'	102	entropy_kinematic_ankle'
27	approxEnt_kinetic'	103	carpiklik_kinematic_ankle'
28	apen_kinetic'	104	basiklik_kinematic_ankle'
29	fast_apen_kinetic'	105	kolmogorov_kinematic_ankle'
30	covariance_kinetic'	106	Higuc_kinematic_ankle'
31	dfa_kinetic'	107	Katz_kinematic_ankle'
32	kinematic_hip_HeelStrike'	108	SampEnt3_kinematic_ankle'
33	kinematic_hip_ToeOff'	109	multiscaSmplEnt_kinematic_ankle'
34	kinematic_hip_MaxFlexion'	110	HarmMean_kinematic_ankle'
35	kinematic_hip_MaxExtension'	111	RMS_kinematic_ankle'
36	min_kinematic_hip'	112	Variance_kinematic_ankle'
37	max_kinematic_hip'	113	trim_kinematic_ankle'

38	ort_kinematic_hip'	114	Mink_kinematic_ankle'
39	medyan_kinematic_hip'	115	Maxk_kinematic_ankle'
40	std_kinematic_hip'	116	Coefficient_Variation_kinematic_ankle'
41	aralik_kinematic_hip'	117	approxEnt_kinematic_ankle'
42	entropy_kinematic_hip'	118	apen_kinematic_ankle'
43	carpiklik_kinematic_hip'	119	fast_apen_kinematic_ankle'
44	basiklik_kinematic_hip'	120	covariance_kinematic_ankle'
45	kolmogorov_kinematic_hip'	121	dfa_kinematic_ankle'
46	Higuc_kinematic_hip'	122	basic_StrideTime'
47	Katz_kinematic_hip'	123	basic_Cadance'
48	SampEnt3_kinematic_hip'	124	basic_StepLenght'
49	multiscaSmplEnt_kinematic_hip'	125	basic_WalkingSpeed'
50	HarmMean_kinematic_hip'	126	min_basic'
51	RMS_kinematic_hip'	127	max_basic'
52	Variance_kinematic_hip'	128	ort_basic'
53	trim_kinematic_hip'	129	medyan_basic'
54	Mink_kinematic_hip'	130	std_basic'
55	Maxk_kinematic_hip'	131	aralik_basic'
56	Coefficient_Variation_kinematic_hip'	132	entropy_basic'
57	approxEnt_kinematic_hip'	133	carpiklik_basic'
58	apen_kinematic_hip'	134	basiklik_basic'
59	fast_apen_kinematic_hip'	135	kolmogorov_basic'
60	covariance_kinematic_hip'	136	Higuc_basic'
61	dfa_kinematic_hip'	137	Katz_basic'
62	kinematic_knee_HeelStrike'	138	SampEnt3_basic'
63	kinematic_knee_ToeOff'	139	multiscaSmplEnt_basic'
64	kinematic_knee_MaxFlexion'	140	HarmMean_basic'
65	kinematic_knee_MaxExtension'	141	RMS_basic'
66	min_kinematic_knee'	142	Variance_basic'
67	max_kinematic_knee'	143	trim_basic'
68	ort_kinematic_knee'	144	Mink_basic'
69	medyan_kinematic_knee'	145	Maxk_basic'
70	std_kinematic_knee'	146	Coefficient_Variation_basic'
71	aralik_kinematic_knee'	147	approxEnt_basic'
72	entropy_kinematic_knee'	148	apen_basic'
73	carpiklik_kinematic_knee'	149	fast_apen_basic'
74	basiklik_kinematic_knee'	150	covariance_basic'
75	kolmogorov_kinematic_knee'	151	dfa_basic'
76	Higuc_kinematic_knee'	152	VKI'

Çizelge 3.2’de gösterilmiş olan özniteliklerin 1-31 arası yürüyüşün kinetik verileri, 32-61 arası yürüyüşte kinematik kalça açısı, 62-91 arası yürüyüşte kinematik diz açısı, 92-121 arası yürüyüşte kinematik ayak bileği açısı, 122-151 arası yürüyüşün temel verileri, 152 ise hastanın vücut kitle indeksi verisidir. Temel, kinetik ve kinematik veri gruplarına ayrı ayrı uygulanan öznitelik çıkarım yöntemleri aşağıda başlıklar halinde verilmiştir.

3.3.1. Minimum Değer

Her bir kişinin seçilen verileri içindeki en küçük değerinin hesaplandığı özneliktir. MATLAB’da bu değer hesaplanması için “*min*” komutu kullanılmıştır.

3.3.2. Maximum Değer

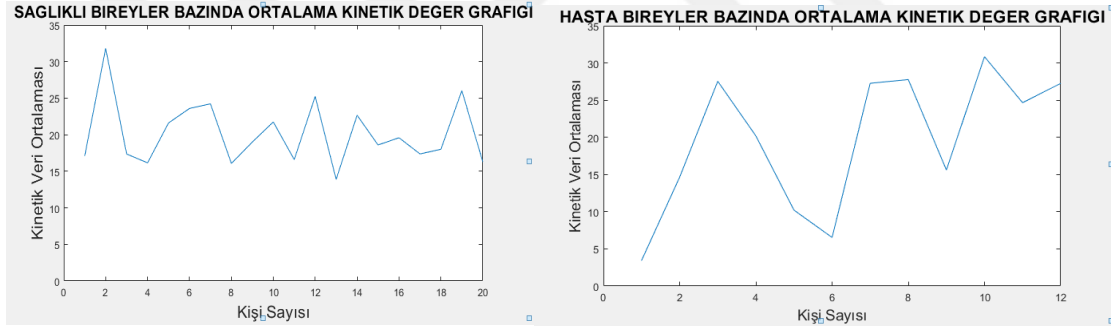
Her bir kişinin seçilen verileri içindeki en büyük değerinin hesaplandığı özneliktir. MATLAB’da bu değer hesaplanması için “*max*” komutu kullanılmıştır.

3.3.3. Ortalama Değer

Her bir kişinin seçilen verileri içindeki ortalama değerinin hesaplandığı özneliktir. MATLAB’da bu değer hesaplanması için “*mean*” komutu kullanılmıştır.

$$A_{ort} = \frac{1}{N} \sum_{i=1}^N A_i$$

Denklem 3.1.



Şekil 3.2. Sırasıyla sağlıklı (a) ve hasta (b) bireylerin kinetik verilerinin ortalama değerleri grafiksel olarak verilmiştir.

Örneğin Şekil 3.2’de sırasıyla sağlıklı ve hastaların birey bazlı kinetik verilerinin ortalama değerleri grafiksel olarak verilmiştir. İki grafik karşılaştırıldığında sağlıklı bireylerde ortalama kinetik değerler yaklaşık 15-30 aralığında iken, Parkinson hastalarında yaklaşık 5-30 aralığındadır. Grafikten de anlaşılacağı üzere ortalama değer gibi özneliklerin hesaplanması ve sınıflandırmada kullanılması anlamlı sonuçların elde edilebilmesi açısından değerlidir.

3.3.4. Medyan (Orta) Değer

Seçilen satırdaki veya sütundaki değerleri küçükten büyüğe doğru sıralayarak ortadaki değer hesaplanır. MATLAB’da bu değer için “*median*” komutu kullanılmıştır. N elemanlı bir kümenin (elemanlar sıralı olmak şartıyla) ortanca değeri $(N+1)/2$ formülü ile hesaplanır.

3.3.5. Standart sapma

Standart sapma, bir verinin veri setinin ortalamasına göre dağılımının ölçülmesidir ve varyansın kareköküdür. Yani, her bir veri noktasının ortalamaya göre ne kadar saptığının belirlenmesidir. Verinin değeri ortalama değerden ne kadar farklıysa standart sapma da o kadar yüksek olur. MATLAB’da bu değer için “*std*” komutu kullanılmıştır.

$$Std = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |A_i - A_{ort}|^2} \quad \text{Denklem 3.2.}$$

3.3.6. Değer Aralığı

Bir matrisin en büyük ve en küçük değeri arasındaki farktır. MATLAB’da bu değer için “*range*” komutu kullanılmıştır.

3.3.7. Entropi

Belirsizliğin ölçüsü anlamına gelen entropi, bir sistemindeki düzensizlik ve rastgelelik hakkında sayısal bilgi verir ve istatistik ve veri madenciliği alanında sıklıkla kullanılır. Entropi değerinin yüksek olması o veri grubunda belirsizliğin fazla olması anlamına gelmektedir. MATLAB’da bu değer için “*entropy*” komutu kullanılmıştır.

3.3.8. Çarpıklık Değeri

Çarpıklık (Skewness), normal dağılımını test eden ölçütlerinden birisidir. Bir dağılımın simetriklik derecesini ölçer. Eğer dağılım simetrik ise çarpıklık değeri sıfırdır. MATLAB’da bu değer için “*skewness*” komutu kullanılmıştır.

3.3.9. Basıklık Deęeri

Basıklık (Kurtosis), 1905 yılında Pearson tarafından açıklanmış bir kavramdır. Deęer dağılımının ortasında çok fazla veya çok az örneęin bulunma derecesini gösteren bir deęerdir. Eęer dağılım eşit ise yani belli deęerlerde yığılmalar yok ise basıklık deęeri sıfırdır. MATLAB’da bu deęerin hesaplanması için “*kurtosis*” komutu kullanılmıştır.

3.3.10. Kolmogorov

Verilerin normal dağılıma sahip olup olmadığını incelemek amacıyla uygulanan normallik testlerinden biridir. Kolmogorov, istatistik alanında en çok bilinen ve kullanılan Ki-Kare normallik testine benzer bir testtir. MATLAB’da Stephen Faul isimli bir araştırmacı tarafından yazılan bir “*kolmogorov*” fonksiyonu ile hesaplanmıştır (Faul, 2021).

3.3.11. Higuchi Fraktal Boyutu (Higuchi Fractal Dimension)

Fraktal boyut hesaplamaları, doğrusal olmayan sinyallerin karmaşıklıklarının hesaplanmasında kullanılmaktadır. MATLAB’da Jesús Monge-Álvarez isimli bir araştırmacı tarafından yazılan bir “*Higuchi_FD*” fonksiyonu ile hesaplanmıştır (Monge-Álvarez, 2021).

3.3.12. Katz Fraktal Boyutu (Katz Fractal Dimension)

Higuchi Fraktal boyut hesaplamasından farklı olarak fraktal boyutun logaritmik olarak hesaplandığı bir yöntemdir. Matrisin ardışık noktaları arasındaki Öklid mesafelerinin toplamı ve ortalaması ve matrisin ilk noktası ile dięer herhangi bir noktası arasındaki maksimum mesafe hesaplanarak logaritmik işlemlerle yapılan bir hesaplamadır. MATLAB’da Jesús Monge-Álvarez isimli bir araştırmacı tarafından yazılan bir “*Katz_FD*” fonksiyonu ile hesaplanmıştır (Monge-Álvarez, 2021).

3.3.13. Sample entropi

Entropi deęeri karmaşık biyolojik işaretlerin incelenmesinde kimi zaman yetersiz kalabilmektedir. Bu sebeple farklı entropi yaklaşımları geliştirilmiştir. Bunlardan biri de

sample entropidir ve bilinen olasılıksal karaktere sahip rasgele sayı kümelerinin analizinde kullanılmaktadır (Richman & Moorman, 2000).

Bu çalışmada sample entropi değerinin hesaplanmasında MATLAB’da UNO Biomechanics araştırmacıları tarafından yazılan bir “*Samp_Ent*” fonksiyonu kullanılmıştır (UNO Biomechanics, 2021).

3.3.14. Multiscale Sample Entropi (MSE-Çok Ölçekli Entropi)

Biyolojik sinyallerin karmaşıklığını ölçmede kullanılan entropi vb. geleneksel yaklaşımlar, zaman serilerinde bulunan çoklu zaman ölçeklerini hesaba katmamaktadır. Bu sebeple yalnızca entropi kullanımı, sağlık ve hastalık durumlarında elde edilen gerçek verilere uygulandığında çelişkili bulgular elde edilebilir. Bu çelişkilerin giderilmesi için zaman serilerinde çoklu zaman ölçeklerini de dikkate alacak çok ölçekli entropi hesaplayabilen yaklaşımlar geliştirilmiştir (Costa, Goldberger & Peng, 2005).

Bu çalışmada MSE değerinin hesaplanmasında MATLAB’da John Malik isimli bir araştırmacı tarafından yazılan ve paylaşılan bir “*multiscaleSampleEntropy*” fonksiyonu kullanılmıştır (Malik, 2021).

3.3.15. Harmonik Ortalama Değeri

Aykırı değerlerden az etkilenen bir ortalama türüdür. Harmonik ortalama, matris elemanlarının terslerinin aritmetik ortalamasının tersi olarak ifade edilebilir. MATLAB’da bu değerın hesaplanması için “*harmmean*” komutu kullanılmıştır.

$$A_{harm} = \frac{n}{\sum_{i=1}^n \frac{1}{A_i}} \quad \text{Denklem 3.3.}$$

3.3.16. RMS (Root Mean Square-Karesel Ortalamanın Karekökü) Değeri

RMS (Karesel ortalamanın karekökü), bir matrisin büyüklüğünün ölçülmesi için kullanılmaktadır. Herhangi bir dalganın ya da matrisin RMS hesaplanırken, sinyal boyunca kaydedilen genlik değerlerinin karelerinin ortalaması alınır (Denklem 3.4). MATLAB’da bu değerın hesaplanması için “*rms*” komutu kullanılmıştır.

$$X_{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^N |X_n|^2}$$
 Denklem 3.4.

3.3.17. Varyans

Varyans değeri, bir matristeki bütün değerlerin ortalamadan uzaklıklarının karelerinin ortalaması ile hesaplanan istatistiksel bir değerdir. MATLAB’da bu değer için “*var*” komutu kullanılmıştır.

$$V = \frac{1}{N} \sum_{i=1}^N |A_i - A_{ort}|^2$$
 Denklem 3.5.

3.3.18. Trim Ortalama Değeri (Kırpılmış Ortalama Değer)

Aykırı değerler hariç olmak üzere bir matrisin ortalama değerinin hesaplandığı yöntemdir. İstatistik, sinyal işleme ve makine öğrenmesi hesaplamalarında sıklıkla kullanılmaktadır. MATLAB’da bu değer için “*trimmean*” komutu kullanılmıştır.

3.3.19. Min-k Değeri

Bir matrisin en küçük k adet değerinin bulunmasında kullanılmaktadır. Bu çalışmada min değeri ayrıca hesaplanıp kullanılmıştır. Ancak k=5 seçilerek matrislerin en küçük 5 değeri bulunmuş ve bu en küçük 5 değerinin ortalaması kullanılmıştır. MATLAB’da bu değer için “*mink*” komutu kullanılmıştır. K=5 için elde edilen 5 en küçük değer “*mean*” komutu ile ortalaması alınarak kullanılmıştır.

3.3.20. Max-k Değeri

Bir matrisin en büyük k adet değerinin bulunmasında kullanılmaktadır. Bu çalışmada max değeri ayrıca hesaplanıp kullanılmıştır. Ancak k=5 seçilerek matrislerin en büyük 5 değeri bulunmuş ve bu en büyük 5 değerinin ortalaması kullanılmıştır. MATLAB’da bu değer için “*maxk*” komutu kullanılmıştır. K=5 için elde edilen 5 en büyük değer “*mean*” komutu ile ortalaması alınarak kullanılmıştır.

3.3.21. Varyasyon Katsayısı (Coefficient Varyasyon)

Standart sapma dağılımının yaygınlığının hesaplandığı bir değerdir. Varyasyon katsayısı, istatistiksel yayılma ölçüsü olarak da tanımlanabilir. Standart sapmanın, ortalama değere oranı ile hesaplanmaktadır.

$$CV = \frac{std}{ort} = \frac{\sqrt{\frac{1}{N-1} \sum_{i=1}^N |A_i - A_{ort}|^2}}{\frac{1}{N} \sum_{i=1}^N A_i} \quad \text{Denklem 3.6.}$$

3.3.22. Approximate Entropi

Doğrusal olmayan serilerin düzenlilik ölçütü olarak tanımlanabilir. Tıpkı MSE gibi, karmaşık biyolojik verilerde yalnızca entropi hesaplamasının yeterli olmadığı hallerde kullanılan bir entropi türüdür (Costa vd., 2005). Esasen bir zaman serisindeki dalgalanmaların tahmin edilemezliğini hesaplar. Bu çalışmada MATLAB’da Approximate entropi hesabı için iki farklı fonksiyon kullanılmıştır. Birincisi MATLAB’ın kütüphanesinde bulunan “*approximateEntropy*” kodu, ikincisi ise “*approx_entropy*” komutu ile Avinash Parnandi tarafından paylaşılan fonksiyondur (Parnandi, 2021).

3.3.23. Fast Approximate Entropi

Fast Approximate Entropy başlığı ile Kijoon Lee tarafından paylaşılan “*ApEn*” komutlu fonksiyondur. Approximate Entropinin daha hızlı ve verimli şekilde hesaplandığı iddia edilmektedir (Lee, 2021).

3.3.24. Kovaryans (Covariance)

İki farklı değişkenin birlikteki değişkenlikleri olarak ifade edilmektedir. MATLAB’da “*cov*” komutu kullanılarak hesaplanmıştır.

$$cov(A, B) = \frac{1}{N-1} \sum_{i=1}^N (A_i - \mu_A) * (B_i - \mu_B) \quad \text{Denklem 3.7.}$$

3.3.25. DFA (Detrended Fluctuation Analysis-Eğilimden Arındırılmış Dalgalanma Analizi)

Çoğu biyolojik sinyal türü durağan değildir. Bu nedenle, DFA gibi trendi giderilmiş zaman serilerinin ortalama karekök dalgalanmasını hesaplayan yöntemlerin uygulanması denenmiştir. Bu yöntem son yıllarda çok çeşitli simüle edilmiş ve fizyolojik zaman serilerine de uygulanan bir yöntemdir (Acharya, Lim & Joseph, 2002). MATLAB’da bu özneteliğin çıkarılması için Martin Magris tarafından yayınlanmış “DFA_fun” fonksiyonu kullanılmıştır (Magris, 2021).

3.4. Normalizasyon Yöntemleri

Eldeki verilerin belirlenen bir aralığa sıkıştırılması normalizasyon olarak tanımlanabilir. Farklı değerlere sahip veriler seçilen bir değer aralığına indirilir (Silahtaroglu, 2016).

Normalizasyonda amaç; veri tabanında veri tekrarının, veri kaybının ya da veri yetersizliğinin önlemektir. Normalizasyon uygulamaları ile veri bütünlüğü sağlanır, gereksiz veri tekrarını önlenir, verilerde meydana gelebilecek bozulmalar önlenir, farklı yöntemlerin uygulanmasında dahi veri tabanı tutarlı olarak çalışır ve sistem performansı artar, daha hızlı ve verimli sistemlerin tasarlanması mümkün olur.

İstatistiksel çalışmalarda ve makine öğrenmesi çalışmalarında kullanılan farklı normalizasyon yöntemleri vardır. Bunlar; min-maks normalizasyon, sıfır ortalama normalizasyon ve ondalıklı normalizasyondur.

Bu tez çalışmasında makine öğrenmesi algoritmalarında en çok kullanılan min-maks normalizasyon ve sıfır ortalama normalizasyon uygulanmıştır. Ayrıca normalizasyon yöntemlerinin başarıya etkisinin belirlenebilmesi için normalizasyon yapılmadan da sınıflandırma sonuçları elde edilmiştir. İlerleyen bölümlerde normalizasyon yapılmadan, min-maks normalizasyon ile ve sıfır ortalama normalizasyon ile elde edilen sonuçların karşılaştırılması ve yorumları bulunmaktadır.

3.4.1. Sıfır-Ortalama Normalizasyon

Sıklıkla kullanılan normalizasyon yöntemlerinden biridir. Bu yöntemde veri ortalama değeri ve standart sapma değerlerinin oranları hesaplanarak yeni bir değeri elde edilmektedir.

$$A_{normalize} = \frac{A - A_{ort}}{std} \quad \text{Denklem 3.8.}$$

3.4.2. Min-Maks Normalizasyon

Verilerin doğrusal olarak normalize edildiği yöntemdir. Bir veri kümesinin 0-1 Aralığına sıkıştırılmasında kullanılan bir yöntemdir.

Min: seçilen veri kümesindeki en küçük değeri, Maks: seçilen veri kümesindeki en büyük değeri.

$$A_{normalize} = \frac{A - Min}{Maks - Min} \quad \text{Denklem 3.9.}$$

Ancak bu çalışmada olduğu gibi değeri aralığının [0,1] olarak seçilmediği durumlarda bu formül farklılık göstermektedir (Denklem 3.10).

$$A_{normalize} = \frac{A - Min}{Maks - Min} (YeniMaks - YeniMin) + (YeniMin) \quad \text{Denklem 3.10.}$$

Bu tez çalışmasında Min-Maks normalizasyon yöntemi uygulanırken değeri aralığı [-1,+1] olarak seçilmiştir. Yani bu uygulamada formülde bulunan YeniMaks=+1, YeniMin=-1 olarak kullanılmıştır.

3.5. Öznitelik Seçme Yöntemleri

Makine öğrenmesi algoritmalarında daha az öznitelikle daha yüksek hıza ve performansa ulaşmak için öznitelik seçimi yapılmaktadır. Öznitelik seçiminde amaç daha az sayıda öznitelik kullanılarak daha hızlı ve verimli bir algoritma oluşturabilmektir. Yani sonucun elde edilmesinde en belirleyici olan özniteliklerin seçilmesidir. Belirleyici öznitelikler seçilip sınıflandırmada kullanıldığında en az sayıda öznitelikle hızlı ve başarı oranı artmış sınıflandırıcılar tasarlanabilmektedir.

Öznitelik seçimi aynı zamanda bilgi kazanımı sağlamayan, sistem başarısını düşüren ya da sınıflandırmaya katkısı olmayan veya az olan özniteliklerin sınıflandırıcılarda kullanılmamasını da sağlayarak, sınıflandırma hızının artırılmasını sağlamaktadır.

Ayrıca makine öğrenmesi yöntemlerinde öznitelik sayısı arttıkça aşırı öğrenme (ezber) ihtimali de artabilmektedir. Aşırı öğrenmenin engellenmesi de yine öznitelik seçimi ile gerçekleştirilebilmektedir.

Öznitelik seçimi için farklı yöntemler kullanılmaktadır. Bu yöntemlere örnek olarak LASSO, mRmR (minimum redundancy maximum relevance: minimum artıklık maksimum ilişki), RELIEF, ardışık ileri (SFS) ve geri yönde seçim (SBS) ve rastgele orman (RO) gösterilebilir.

Bu tez çalışmasında öznitelik seçim yöntemlerinden üçü kullanılmış ve kullanılan sınıflandırma algoritmalarında sonuçlar ayrı ayrı denenerek yorumlanmıştır. Kullanılan öznitelik seçme yöntemleri:

1. LASSO (Least Absolute Shrinkage and Selection Operator- En Küçük Mutlak Daralma ve Seçme Operatörü)
2. mRmR (Minimum Redundancy Maximum Relevance-Minimum Artıklık Maksimum İlişki)
3. RELIEF

3.5.1. LASSO (Least Absolute Shrinkage and Selection Operator- En Küçük Mutlak Daralma ve Seçme Operatörü)

İlk olarak 1996 yılında Robert Tibshirani tarafından formüle edilmiştir. LASSO öznitelik seçim yöntemi hem öznitelik seçimi yaparken hem de regresyonu aynı anda yapması açısından avantajlı bir yöntemdir. LASSO, parametrelerin mutlak değerlerinin toplamına bir kısıtlama getirir, toplamın sabit bir üst sınır değerinden küçük olması

gereklidir. Bunu yapmak için yöntem, bazı parametreleri sıfıra indiren değişkenlerin katsayılarını sıfır yaparak cezalandırıcı bir işlem uygular. Öznitelik seçim sürecinde, küçültme işleminden sonra hala sıfır olmayan bir katsayıya sahip olan değişkenler seçilerek modelin parçası olarak tayin edilir. Bu işlemin amacı tahmin hatasını en aza indirmektir (Fonti & Belitser, 2017).

LASSO yöntemini kullanmak, çok iyi bir tahmin doğruluğu sağlayabilmesi açısından avantajdır. Çünkü katsayıları küçültmek ve kaldırmak, sapmada önemli bir artış olmadan varyansı azaltır. Bu özellik hem az sayıda veri içeren veri kümelerinde hem de çok sayıda özniteliğe sahip veri kümelerinde kullanma olanağı sunar (Fonti & Belitser, 2017).

LASSO aynı zamanda aşırı öğrenmeyi (overfitting) önlemenin yanı sıra, önemsiz ya da az önemli özniteliklerin katsayılarını sıfıra indirir. Bu durum sınıflandırıcı başarısını düşüren özniteliklerin elenmesini sağlayarak doğruluk oranı daha yüksek sonuçlar elde etmeyi sağlar.

3.5.2. Minimum Artıklık Maksimum İlişki (mRmR)

mRmR, başarı oranı ve çalışma hızı yüksek olan entropi tabanlı bir öznitelik seçme yöntemidir. Entropi temelli bir yöntem olup hem sınıflandırmaya katkısı en çok olan öznitelikleri seçerken hem de öznitelikler arası artıklığı en aza indirerek filtreleme işlemi yapar.

$$I(X, Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log\left(\frac{p(x, y)}{p_1(x)p_2(y)}\right) \quad \text{Denklem 3.11.}$$

Denklem 3.11. ifadesinde $p(x, y)$: X ve Y'nin birleştirilmiş olasılık dağılımını ifade eden fonksiyondur, $p_1(x)$: X özniteliğinin marjinal olasılık dağılım fonksiyonu ve $p_2(y)$: Y özniteliğinin marjinal olasılık dağılım fonksiyonunu ifade eder.

S:seçilecek olan özniteliklerin kümesi ise, $|S|$:S kümesinin eleman sayısını ifade eder. Denklem 3.12'de minimum artıklık (W) Denklem 3.13'te maximum ilişki (V) formülleri verilmiştir. Burada F_i , ayrık rastlantı değişkenlerini, $I(F_i, F_j)$ ise $(i, j=1, 2, 3, \dots, d, d$: veri kümesi boyutu olmak üzere) i ve j arası ortak bilgiyi ifade eder.

$$W = \frac{1}{|S|^2} \sum_{F_i, F_j \in S} I(F_i, F_j) \quad \text{Denklem 3.12.}$$

$$V = \frac{1}{|S|} \sum_{F_i \in S} I(F_i, H) \quad \text{Denklem 3.13.}$$

Yukarıdaki denklemlere ek olarak V ve W değerlerinin kombinasyonları max(V-W) ve max(V/W) şeklinde elde edilerek optimize edilir. Sonuçta mRmR yöntemi iki farklı kriter belirleyerek öznitelik seçimi yapar. Bu kriterler ortak bilgi farkı (MID) ve ortak bilgi oranı (MIQ), Denklem 3.14. ve Denklem 3.15'te verilmiştir (Gülgezen vd.).

$$MID = \max_{F_i, \Omega_S} [I(F_i, H) - \frac{1}{|S|} \sum_{F_j \in S} I(F_i, F_j)] \quad \text{Denklem 3.14.}$$

$$MIQ = \max_{F_i, \Omega_S} [I(F_i, H) / \frac{1}{|S|} \sum_{F_j \in S} I(F_i, F_j)] \quad \text{Denklem 3.15.}$$

3.5.3. RELIEF

Relief öznitelik seçme yöntemi, 1992 yılında Kira ve arkadaşları tarafından geliştirilmiştir ve aralarındaki ilişkiye göre öznitelikleri ağırlıklandırır. Bu yöntem özellikle iki sınıf içeren veri setleri için ciddi anlamda başarılı sonuç vermektedir. Ancak iki sınıftan fazla sınıfa sahip veri setleri için uygun değildir. Bu tez çalışmasında Relief yönteminin kullanılmasının en önemli sebebi budur. Yöntemin ilk adımında tüm özniteliklerin ağırlıkları 0 atanır. Sonrasında her adımda veri setinden rastgele bir veri seçilir. Seçilen bu veri ile aynı sınıfta bulunan en yakın k adet veri bulunur ($k=(\text{SınıfSayısı}-1)$ olmak üzere). Devamında her sınıfa ait en yakın veri bulunur. Ardından her bir özniteliğin ağırlığı bu veriler kullanılarak güncellenir. Son olarak belirlenen koşulu sağlamayan özellikler veri setinden çıkarılarak yeni veri seti elde edilir (Kaynar, Arslan, Görmez & Demirkoparan, 2017).

$$W(x^a) = W(x^a) - \frac{\sum_{j=1}^k \text{uzaklık}(A, R_i, H_j)}{m*k} + \frac{\sum_{C \neq \text{sınıf}(R_i) [\frac{P(C)}{1-P(\text{sınıf}(R_i))}] * \sum_{j=1}^k \text{uzaklık}(A, R_i, M_j)}]}{m*k} \quad \text{Denklem 3.16.}$$

Denklem 3.16'te verilen formülde $W(x^a)$: a. özniteliğin ağırlığını, k: sınıf sayısının bir eksiğini, R_i : i. döngüde seçilen veriyi, H_j : seçilen veri ile aynı sınıfta olan j. yakın veriyi, M_j : j. sınıfa ait seçilen veriye en yakın olan veriyi, m: döngü sayısını ifade etmektedir.

$$uzaklık(A, I_1, I_2) = \begin{cases} 0, & I_1 = I_2 \\ 1, & I_1 \neq I_2 \end{cases}, uzaklık(A, I_1, I_2) = |I_1 - I_2| * \frac{1}{\max(A) - \min(A)} \text{ Denklem 3.17.}$$

3.6. Sınıflandırma Yöntemleri

Hangi sınıfa ait olduğu bilinmeyen verilerin, öznitelikler kullanılarak sınıflarının belirlenmesi işlemidir. Makine öğrenmesi algoritmaları kullanılarak yapılan sınıflandırmalar, karmaşık ve çok sayıda verinin özniteliklerinin analiz edilmesini, analizlerden anlamlı sonuçlar çıkarılarak verilerin sınıflandırılmasını hızlı, güvenilir ve yüksek başarı ile gerçekleştirir.

Makine öğrenmesinde sınıflandırma, denetimli öğrenme ile gerçekleştirilir. Denetimli öğrenme, etiketlenmiş yani sınıf bilgileri belirlenmiş eğitim verilerinden bir fonksiyon oluşturularak makinenin öğreniminin sağlanması ve test verilerinin de bu fonksiyonda çalıştırılarak tahminin doğruluk analizinin yapılmasını ifade eder.

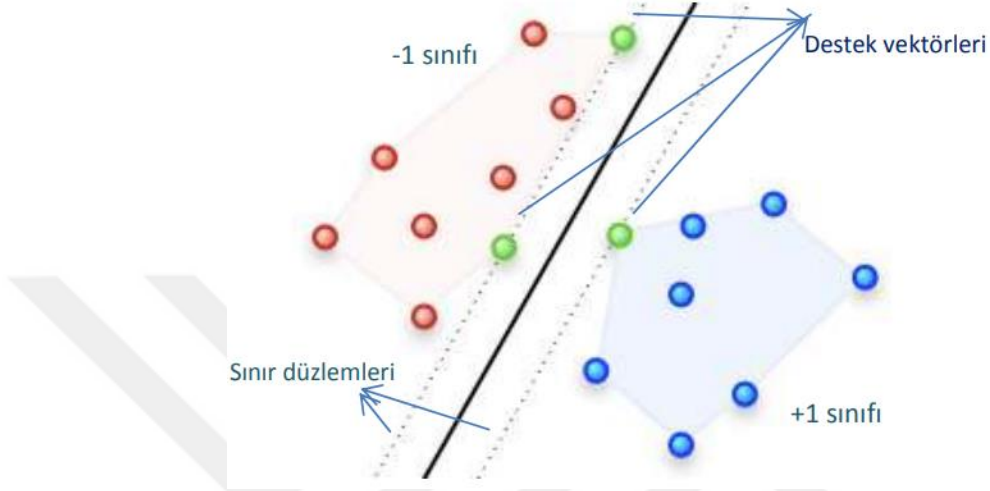
Denetimli öğrenme algoritmalarına örnek olarak: KNN, Doğrusal Regresyon, Lojistik regresyon, Rastgele Orman, Gradyan Güçlendirilmiş Ağaçlar, SVM, Naive Bayes, Nöral ağlar ve Karar ağaçları verilebilir.

Bu tez çalışmasında sınıflandırma algoritması olarak; KNN, SVM ve Naive Bayes sınıflandırma yöntemleri kullanılmıştır.

3.6.1. SVM (Support Vector Machines/Destek Vektör Makineleri)

SVM sınıflandırıcısı temelde verileri bir düzlem ya da bir hiper düzlem ile ayırarak sınıflandırma işlemidir. Yani esasen amaç, iki sınıfa ait olan verilerin ayrımında en uygun düzlemi veya hiper düzlemi belirleyebilmektir. Burada verilerin doğrusal olarak ayrılabilme ya da ayrılamama durumları söz konusudur. Veriler eğer doğrusal olarak ayrılabilir ise, veri setleri ile aynı boyutta bir düzlem oluşturulabilir. Doğrusal olarak ayrımın yapıldığı durumlarda sonsuz sayıda doğru çizilerek bu doğrular içinden marjini en yüksek yapabilecek doğru seçilir. Ancak eğer veriler doğrusal olarak ayrılamıyor ise veri setleri boyutundan daha büyük boyutta bir hiper düzlem oluşturulur ve ayırım yapılır. Doğrusal olarak ayrımı yapılamayan verilerin olduğu durumda SVM, veriyi çok daha

yüksek boyutlu bir uzaya taşır ve burada verileri sınıflandırmak için optimum bir hiper düzlem oluşturur (Akşehirli, Ankaralı, Aydın & Saraçlı, 2013). Şekil 3.3'te SVM sınıflandırıcısının iki sınıfı ayırmak için oluşturduğu temsili bir düzlem gösterilmiştir.



Şekil 3.3. SVM sınıflandırıcısında iki sınıfın ayrılmasında oluşturulmuş bir hiper düzlem görseli (Ayhan & Erdoğan, 2014).

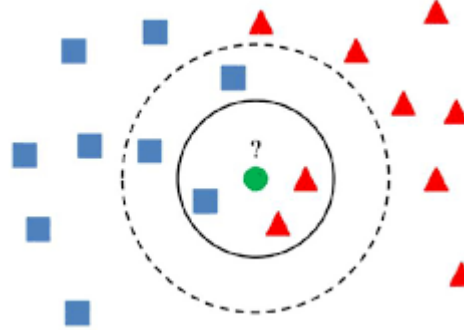
Özellikle karmaşık verilerin bulunduğu ve sınıf sayısının az olduğu sınıflandırmalar için uygun bir yöntemdir. Bu sebeple bu çalışmada SVM sınıflandırma yöntemi seçilmiştir.

3.6.2. KNN (K-Nearest Neighbors/K-En Yakın Komşular)

KNN sınıflandırma algoritması, uygulanabilirliğinin basit olması, analitik olarak izlenebilirliğinin olması, öğrenme sürecinin hızlı ve yüksek başarı oranına sahip olması sebepleri ile sınıflandırma algoritmaları arasında yaygın olarak kullanılmaktadır. Ayrıca farklı boyutta veri setleri üzerinde de uygulanabilmekte ve gürültülü verilerde dahi yüksek doğruluğa sahip sonuçlar elde edilebilmektedir. Tüm bu avantajlarının yanında bellek gereksiniminin fazla olması, veri ve öznitelik sayısı ve boyutu arttıkça işlem yük ve maliyetinin de artması, performansın k komşu sayısı, uzaklık ölçütü ve öznitelik miktarı gibi parametrelere göre etkilenmesi gibi dezavantajları da bulunmaktadır (Taşcı & Onan, 2016).

KNN sınıflandırıcıda etiketlenmiş veriler n boyutlu bir uzaya taşınır. Sınıfı bilinmeyen yeni test verisi uzayda kendisine en yakın k sayıdaki komşunun oluşturduğu sınıfa dâhil olduğu tahmin edilir. Bu tahminin oluşmasında sınıflar arasında test verisine en çok komşu olan ve en yakın verilerin bulunduğu sınıf belirleyicidir (Mete, Çakır, Bayat, Duru & Duru, 2020).

KNN sınıflandırıcıda komşuluk sayısı ve uzaklık bağıntısı sınıflandırıcı başarısını doğrudan etkilemektedir. Sınıflandırıcıda komşuluk sayısı k değeri ile gösterilir ve tek sayı olması tercih edilir. Çünkü çift sayı olduğunda eşit sayıda komşuluk olması durumu hangi sınıfın belirleneceği hususunu karmaşıklaştırır. Uzaklık bağıntıları ise Euclidean (Öklid), Cityblock, Chebychev, Korelasyon, Cosine, Jaccard, Hamming, Minkowski, Seucclidean ve Spearman şeklindedir. Bu tez çalışmasında KNN sınıflandırmada k değeri olarak 3, 5, 7 ve 9 değerleri; uzaklık bağıntısı olarak da en çok kullanılan Öklid, Cityblock, Chebychev ve Minkowski kullanılmıştır.



Şekil 3.4. KNN sınıflandırıcı modeli temsili şeması (Altunkaynak, Başakın & Kartal, 2020).

Şekil 3.4'te iki kırmızı ve mavi renkli şekiller iki farklı sınıf oluşturmaktadır. Yeşil renkli örneğin hangi sınıfa ait olduğunun tahmininde k komşuluk değeri dikkate alınacak olur ise; $K=3$ seçildiğinde örnek kırmızı sınıfa dahil edilirken, $K=5$ seçildiğinde örnek mavi sınıfa dahil edilir. $K=3$ için küçük daire göz önüne alınır ve en yakın üç komşuya bakılır. $K=3$ kümesinde en yakın üç komşudan ikisi kırmızı olduğundan örneğin kırmızı sınıfına dâhil edilmesi gereklidir denir. $K=5$ için büyük daire göz önüne alınır ve en yakın beş komşuya bakılır. $K=5$ kümesinde en yakın beş komşudan üçü mavi olduğundan örneğin mavi sınıfına dâhil edilmesi gereklidir denir.

3.6.3 Naive Bayes

Naive Bayes sınıflandırıcı ismini İngiliz matematikçi Thomas Bayes' ten alır ve olasılık temelli bir sınıflandırma yöntemidir. Bu yöntemde elde var olan veriler kullanılarak sınıflar oluşturulur ve sınıfı bilinmeyen yeni bir verinin bu mevcut sınıflara dâhil olma olasılıklarını hesaplar (Silahtaroglu, 2016).

$$P(C_1|x_i) = \frac{P(x_i|C_1)P(C_1)}{P(x_i|C_1)P(h_1)+P(x_i|C_2)P(C_2)+\dots} \quad \text{Denklem 3.18.}$$

Denklem 3.18'de verilen formülde $P(C_1|x_i)$, x_i 'nin C_1 sınıfında olma olasılığını ifade eder. $P(x_i)$, x_i 'nin veri tabanındaki sıklığını; $P(C_1)$ ve $P(C_2)$ de, C_1 ve C_2 'nin sıklığını ifade etmektedir. Denklem 3.18'in paydasındaki toplam, sınıf sayısı kadar artırılır. h_1 burada hipotez olarak ifade edilir ve hipotez sınıf olarak kabul edilir ise formül, Denklem 3.19'deki şekilde sadeleşir.

$$P(C_1|x_i) = \frac{P(x_i|C_1)P(C_1)}{P(x_i)} \quad \text{Denklem 3.19.}$$

3.7. Çapraz Doğrulama Yöntemleri

Çapraz doğrulama, kısıtlı veri setlerinde performansı arttırmaya yönelik olarak veri setini birkaç bölüme ayırma işlemidir. Veri setinde eğitim ve test işlemleri için aynı verilerin kullanılması sistemin aşırı öğrenmesine ve gerçekçi olmayan sonuçların elde edilmesine neden olabilir. Bu sorunun ortadan kaldırılması için veri seti eğitim ve test için ayrı gruplara ayrılır. Eğitimde kullanılan veriler ile testte kullanılan veriler birbirinden farklıdır. Böylelikle algoritma test aşamasında eğitim aşamasında kullandığı verilerden farklı veriler kullanılır. Yani algoritmayı hiç tanımadığı verilerle test eder. Böylelikle daha gerçekçi bir test işlemi yapılmış olur ve daha gerçekçi bir doğruluk oranı elde edilir.

Bu tez çalışmasında çapraz doğrulama yöntemlerinden K katlı çapraz doğrulama (K fold cross validation) ve birini dışarıda bırakma çapraz doğrulama (Leave-one-out cross validation) kullanılmıştır.

3.7.1. Birini Dışarda Bırak Çapraz Doğrulama (Leave-one-out cross validation)

Küçük boyutlu veri setlerinde uygulanan ve sistem performansını arttıran bir yöntemdir. Verilerin biri test amaçlı olarak eğitim dışında tutulur, kalan verilerin tamamı sistemin eğitilmesi amacıyla kullanılır. Eğitim bittiğinde ayrılan test verisi sistem doğruluğunu test etmek amacıyla kullanılır ve bir doğruluk değeri elde edilir. Bu işlem sıralı olarak tüm veriler tek tek test olacak şekilde tekrarlanır ve her tekrarda bir doğruluk değeri elde edilerek kaydedilir. Sonuçta her bir test verisi için kaydedilen doğruluk değerinin ortalaması alınarak genel bir doğruluk değeri elde edilir. Bu değer sınıflandırıcının doğruluk değeridir.

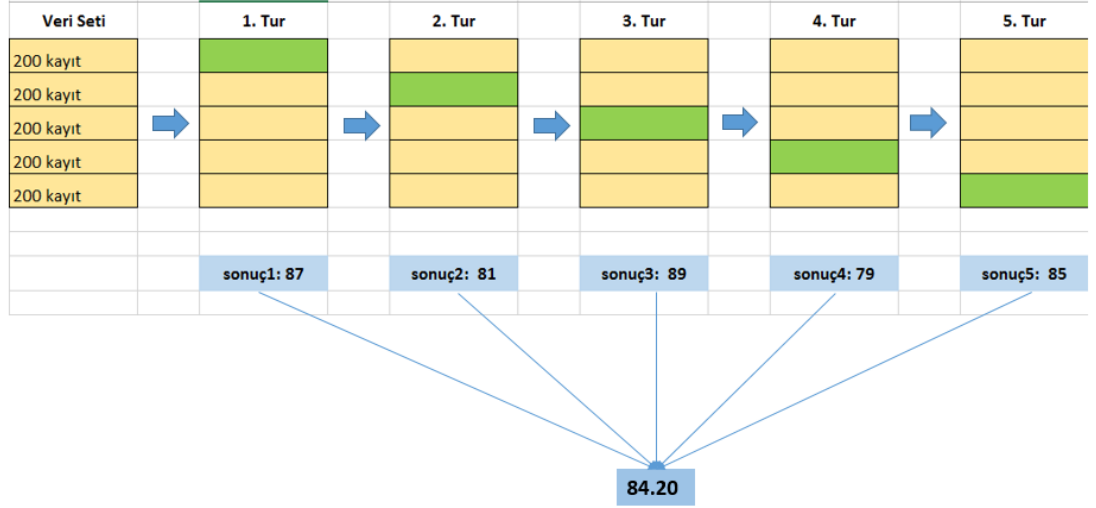
Çizelge 3.3. n adet veriye sahip bir veri setinin birini dışarıda bırakma çapraz doğrulama yöntemi ile çapraz doğrulanması ve sistem performansı hesabı.

1. İterasyon	TEST	...	→Performans 1	<i>Sistem Performansı</i> $= \frac{1}{n} \left(\sum_{i=1}^n Performans_i \right)$
2. İterasyon		TEST	→Performans 2	
...		...	→...	
n. iterasyon		...	TEST	

3.7.2. KFold Çapraz Doğrulama (KFold Cross Validation)

K Fold çapraz doğrulama yönteminde veri seti test ve eğitim kümelerine ayrılmaktadır. Bir K değeri belirlenir ve veri seti eğitim ve test işlemleri için K değeri kadar alt gruba ayrılır. Her alt grup sırasıyla bir kez test kümesi olmak zorundadır. Veri setinde her bir test grubundan artakalan verilen tamamı eğitim grubu olur. Birini dışarıda bırakma çapraz doğrulama yönteminden farklı olarak Kfold çapraz doğrulama yönteminde bir veri değil verilerden oluşan bir grup test işleminde kullanılır. Burada n adet veriye sahip bir veri setinde; n/K test grubunun veri sayısıdır.

Kfold çapraz doğrulama yönteminde de birini dışarıda bırakma çapraz doğrulama yönteminde olduğu gibi her test işleminde bir doğruluk değeri elde edilir ve kaydedilir. Sistem doğruluğu tüm test doğruluklarının aritmetik ortalamasıdır.



Şekil 3.5. Kfold çapraz doğrulama yönteminde test işlemi ve doğruluk hesabı örneği (Veribilimiokulu, 2021)

3.8. Değerlendirme Ölçütleri

Makine öğrenmesi yöntemlerinin başarı ölçümünde farklı yöntemler kullanılmaktadır. Değerlendirme ölçütleri sistemin doğruluğunu, hassasiyetini, özgüllüğünü kısaca performansını ölçmede kullanılır. Değerlendirme ölçütleri sınıflandırma sonucunda doğru ve yanlış tahminlerin yer aldığı karmaşıklık matrisi (confusion matrix) ya da diğer ismi ile hata matrisi üzerinden hesaplanmaktadır. Çizelge 3.4'te karmaşıklık matrisi ve matris içi değerleri gösterilmiştir.

Çizelge 3.4. Sınıflandırıcı sonuçlarının değerlendirme ölçütlerinin hesaplanmasında kullanılan karmaşıklık matrisi

		TAHMİN	
		POZİTİF	NEGATİF
GERÇEK	POZİTİF	Gerçek Pozitif (GP)	Yanlış Negatif (YN)
	NEGATİF	Yanlış Pozitif (YP)	Gerçek Negatif (GN)

Çizelge 3.4'te verilen karmaşıklık matrisi, sınıflara ait verilerin ne kadarının doğru ve ne kadarının yanlış tahmin edildiğini gösterir. Pozitif sınıfa ait bir veri pozitif sınıfta tahmin edilmiş ise bu veri gerçek pozitif (GP), negatif sınıfta tahmin edilmiş ise yanlış

pozitif (YP)'tir. Tam tersi; negatif sınıfa ait bir veri sınıflandırıcı tarafından negatif olarak tahmin edilmiş ise bu veri gerçek negatif (GN), pozitif sınıfta tahmin edilmiş ise yanlış negatif (YN) olarak isimlendirilir. Burada yer alan GP, YP, GN ve YN değerlerinin toplamı veri sayısını verir.

Çizelge 3.4. sonuçlarından yola çıkılarak bu tez çalışmasında doğruluk, hassasiyet, özgülük, kesinlik, hatırlama (geri çağırma), f ölçütü, gmean ve Matthews Korelasyon Katsayısı değerlendirme ölçütleri hesaplanmış ve sonuçların değerlendirilmesinde kullanılmıştır.

3.8.1. Doğruluk (Accuracy)

Sınıflandırma işleminin sonucunda elde edilen doğru sınıflandırılmış veri oranını temsil eder. Doğru sınıfta bulunan veri sayısının tüm veri sayısına oranıdır.

$$\text{Doğruluk} = \frac{GP+GN}{GP+GN+YP+YN} \quad \text{Denklem 3.20.}$$

Bu tez çalışmasında Parkinson hastalarının ve sağlıklı bireylerin ne kadarının doğru tahmin edildiğini ifade etmektedir.

3.8.2. Hassasiyet (Sensitivity)

Pozitif sınıfta tahmin edilmiş pozitif verilerin, tüm gerçek pozitif verilere oranını temsil eder. Bu tez çalışmasında hasta olarak tahmin edilen hasta sayısının toplam hasta sayısına oranını ifade eder.

$$\text{Hassasiyet} = \frac{GP}{GP+YN} \quad \text{Denklem 3.21.}$$

3.8.3. Özgülük (Specificity)

Negatif sınıfta tahmin edilmiş negatif verilerin, tüm gerçek negatif verilere oranıdır. Bu tez çalışmasında sağlıklı olarak tahmin edilen sağlıklı bireylerin, tüm sağlıklı bireylere oranını ifade eder.

$$\text{Özgüllük} = \frac{GN}{GN+YP} \quad \text{Denklem 3.22.}$$

3.8.4. Kesinlik (Precision)

Gerçek değeri pozitif olup, pozitif değere sınıflandırılan verilerin; pozitif değere sınıflandırılan toplam veriye oranıdır. Bu tez çalışmasında hasta olarak tahmin edilen hasta bireylerin, hasta olarak tahmin edilenlerin toplamını ifade eder. Yani hasta olarak tahmin edilenlerin kaçının gerçekten hasta olduğunu ifade eder.

$$Kesinlik = \frac{GP}{GP+YP} \quad \text{Denklem 3.23.}$$

3.8.5. Recall (Hatırlama-Geri Çağırma)

Gerçek değeri pozitif olup, pozitif değere sınıflandırılan verilerin; gerçek değeri pozitif olan tüm verilerin oranıdır. Esasen hassasiyet değeri ile eşittir. Bu tez çalışmasında hasta olarak tahmin edilen hasta sayısının toplam hasta sayısına oranını ifade eder.

$$Recall = \frac{GP}{GP+YN} \quad \text{Denklem 3.24.}$$

3.8.6. F-ölçütü

İki sınıflandırıcı ölçütünün tek bir ölçüt olarak değerlendirilmesini sağlar. F-ölçütü, harmonik ortalama olarak da isimlendirilir. Tek başlarına anlamlı yorumlamalara imkan veremeyecek olan kesinlik ve recall değerlerinin harmonik ortalaması alınarak elde edilen bir ölçüttür.

$$F \text{ ölçütü} = \frac{2 * Kesinlik * Recall}{Kesinlik + Recall} \quad \text{Denklem 3.25.}$$

3.8.7. Gmean

Hassasiyet ve özgüllük değerlerinin çarpımlarının karekökü şeklinde bulunan ölçüttür.

$$gmean = \sqrt{\frac{GP}{GP+YN} * \frac{GN}{GN+YP}} \quad \text{Denklem 3.26.}$$

3.8.8. Matthews Korelasyon Katsayısı (Mcc Değeri)

Matthews Korelasyon Katsayısı, yapılan sınıflandırmadaki tahmin kalitesini gösteren bir ölçüttür. Sınıflar arasında eleman sayısında dengesizlik olması halinde dahi,

sınıflandırma işleminin rastgele yapılıp yapılmadığı ile ilgili bilgi veren bir ölçüttür. Sınıflandırma sonucunun rastgeleliliği sınıflandırmalar arasındaki korelasyon ile hesaplanmaktadır. -1 ila +1 Aralığında bir değer almalıdır. Sınıflandırıcının doğru sınıf tahmini arttıkça mcc değeri +1 değerine yaklaşır, sınıflandırıcının yanlış sınıf tahmininde ise mcc değerinin -1 değerine yaklaşması beklenir. Sınıflandırmanın rastgele yapılması halinde ise mcc değeri 0 civarındadır.

$$mcc \text{ Değeri} = \frac{(GP*GN)-(YN*YP)}{\sqrt{(GP+YN)*(GP+YP)*(GN+YN)*(GN+YP)}}$$

Denklem 3.27.

BÖLÜM 4

DENEYSEL SONUÇLAR

Bu tez çalışmasında Parkinson hastası ve sağlıklı bireylerden elde edilmiş olan yürüyüş verileri incelenerek makine öğrenmesi yöntemleri ile hasta-sağlıklı ayrımı yapabilen bir modelin geliştirilmesi amaçlanmıştır.

Bu çalışmada veri seti olarak Malezya MARA Teknoloji Üniversitesi'nden Prof. Nooritawati Md Tahir ve ekibinin Parkinson hastası ve sağlıklı bireylerden elde ettiği yürüyüş verileri kullanılmıştır. Veriler 12 Parkinson hastası ve 20 sağlıklı bireyin yürüyüşünden elde edilen ve temel, kinetik ve kinematik verileri kapsayan toplamda 21 öznelikten oluşmaktadır.

Sistem oluşturma ve geliştirme kapsamında hâlihazırda var olan özneliklere öznelik çıkarım yöntemleri uygulanarak öznelik sayısı artırılmıştır. Ham verilere eklenen yeni özneliklere farklı normalizasyon yöntemleri uygulanarak sistem performansının artırılması amaçlanmıştır. Elde edilen sınıflandırma sonuçları kaydedilmiştir. Sonrasında öznelik seçme yöntemleri de uygulanarak hangi özneliklerin belirleyici olduğu tespit edilmiş, bu belirleyici özneliklerin sistem performansına katkısı değerlendirilmiştir. Yalnızca belirleyici öznelikler kullanılarak

yeni sınıflandırma sonuçları da ayrıca kaydedilmiştir. Kayıtlı sonuçlar karşılaştırılarak yorumlanmıştır.

Farklı öznitelik seçme algoritmalarından seçilen öznitelikler incelenerek ortak olanlar kaydedilmiş, ortak özniteliklerin sınıflandırıcı performansları da ayrıca analiz edilmiştir.

Yukarıda bahsi geçen işlem adımlarının tamamında MATLAB (The MathWorks, Inc.) uygulaması kullanılmıştır.

4.1. Normalizasyon Yöntemlerinin Etkisi

Bu bölümün ilk aşamasında verilere herhangi bir işlem yapılmadan, ham halleri ile sınıflandırma sonuçları elde edilmiştir. Sonrasında öznitelik çıkarım aşamasına geçilmeden önce ham verilere iki farklı normalizasyon yöntemi uygulanmıştır.

Çizelge 4.1’de SVM, KNN ve Naive Bayes sınıflandırıcılarında çalıştırılan ham verilerin normalizasyon yapılmayan, Min-Maks normalizasyon ve Sıfır ortalama normalizasyon uygulanan kombinasyonlarından elde edilen doğruluk değerleri verilmiştir. Bu tabloda verilen sonuçlar aynı ham veri kümelerine yalnızca farklı normalizasyon tekniklerinin uygulanması ile elde edilmiştir, amaç normalizasyonun sınıflandırma performanslarına etkisini gözlemlemektir.

Çizelge 4.1. Öznitelik çıkarımı öncesi ham veri ile elde edilen sınıflandırıcı doğrulukları ve normalizasyon yöntemlerinin etkisi.

Sınıflandırma Yöntemi	Uygulanan Normalizasyon Yöntemi		
	Normalizasyon Yok	Min-Maks Normalizasyon	Sıfır Ortalama Normalizasyon
SVM	75,00	78,13	62,50
KNN	84,38	<u>87,50</u>	84,38
Naive Bayes	71,88	78,13	62,50

Normalizasyon yöntemleri uygulanarak gereksiz veri tekrarının önüne geçilerek sınıflandırıcı performansının artırılması ve algoritmanın çalışma süresinin kısaltılması amaçlanmıştır.

Çizelge 4.1’de gösterilen sonuçlar sınıflandırma yöntemlerinin mümkün olan tüm parametreleri çalıştırılarak elde edilen sonuçlar içinden en iyi doğruluk oranının elde edildiği sonuçlardır. Çizelge 4.1’e göre; Min-Maks normalizasyon uygulandığında ve KNN sınıflandırıcı algoritması kullanıldığında en yüksek doğruluk oranına ulaşılmıştır. Min-Maks Normalizasyon yönteminin uygulandığı tüm sınıflandırıcılarda, normalizasyon yapılmama durumuna göre doğruluk oranında artış gözlenmiştir. Yani Min-Maks normalizasyon tekniği bu veri setinin sınıflandırılmasında kullanılan bu üç sınıflandırıcıda da performansı artırıcı etki oluşturmuştur. Bu durum bu veri seti üzerinde Min-Maks normalizasyon yönteminin daha başarılı sonuçlar verebileceği fikrini oluşturmuştur. Tablodaki diğer kısımlar incelendiğinde ise Sıfır Ortalama Normalizasyon tekniğinin sınıflandırıcı başarısına olumlu bir katkısının olmadığı görülmektedir. Sıfır Ortalama Normalizasyon tekniği ya sınıflandırıcı başarısını düşürmüştür ya da etkilememiştir.

Örneğin SVM sınıflandırıcısı için 4 farklı Kernel parametresinin tamamı elde edilmiştir. Bu dört farklı parametre: lineer, polinomal, radyal ve sigmoid yaklaşımlarıdır. Çizelge 4.2’de SVM sınıflandırıcı yönteminde farklı normalizasyon yöntemleri ile farklı yaklaşım parametrelerinden elde edilen sonuçlar bulunmaktadır. Çizelge 4.2’de de görüldüğü üzere Sıfır ortalama normalizasyon yöntemi bu veri seti ve sınıflandırma yöntemi için uygun değildir. Normalizasyon yapılmaması durumuna göre, sınıflandırıcı başarısını düşürdüğü görülmektedir. En yüksek doğruluk oranına normalizasyon yapılmayan durumda 2-RBF parametresinde ulaşılmış olmasına rağmen, Min-Maks normalizasyon yöntemi uygulandığında farklı parametrelerden elde edilen doğruluk oranlarının ortalama değeri daha yüksek elde edilmiştir.

Çizelge 4.2. SVM sınıflandırıcıdan elde edilen farklı sonuçlar (Ham veri ile).

	Normalizasyon Yok	Min-Maks Normalizasyon	Sıfır Ortalama Normalizasyon
0=LİNEER	59,38	62,50	NaN
1=POLİNOMAL	68,75	62,50	62,50
2=RADYAL (RBF)	<u>75,00</u>	68,75	37,50
3=ŞİGMOİD	59,38	78,13	43,75
Ortalama doğruluk	65,63	67,97	35,94

KNN sınıflandırıcı yönteminde ise komşuluk sayısı ve uzaklık parametreleri değiştirilerek farklı sonuçlar elde edilmiş ve Çizelge 4.1’de en yüksek doğruluk oranı paylaşılmıştır. KNN sınıflandırıcıda komşuluk sayısı olarak 3, 5, 7 ve 9 değerlerinin tamamı denenmiştir. Uzaklık parametresi olarak da Öklid, Cityblock, Chebychev ve Minkowski uzaklık bağıntıları çalıştırılmıştır. Çizelge 4.3’te KNN sınıflandırıcıda ham veriler ile elde edilen sonuçların detayları bulunmaktadır. En yüksek doğruluk oranına Min-Maks normalizasyonun uygulandığı durumda ulaşılmıştır. Min-Maks normalizasyon KNN sınıflandırıcıda, bu veri setinde, genel olarak sınıflandırma doğruluğunu arttıran bir etken olmuştur. Ancak Sıfır ortalama normalizasyon yöntemi incelendiğinde genel olarak sınıflandırıcı başarısını düşürdüğü gözlemlenmiştir.

Çizelge 4.3. KNN Sınıflandırıcı parametreleri ve ham veriden elde edilen sonuçlar.

	KNN Uzaklık bağıntıları	KOMŞULUK SAYISI			
		3	5	7	9
Normalizasyon Yok	Öklid	78,13	81,25	75,00	62,50
	Cityblock	<u>84,38</u>	<u>84,38</u>	78,13	78,13
	Chebychev	75,00	71,88	75,00	71,88
	Minkowski	78,13	81,25	75,00	62,50
Min-Maks Normalizasyon	Öklid	<u>87,50</u>	78,13	75,00	75,00
	Cityblock	<u>87,50</u>	<u>87,50</u>	81,25	75,00
	Chebychev	81,25	81,25	75,00	68,75
	Minkowski	<u>87,50</u>	78,13	75,00	75,00
Sıfır Ortalama Normalizasyon	Öklid	81,25	71,88	65,63	59,38
	Cityblock	<u>84,38</u>	68,75	71,88	59,38
	Chebychev	59,38	68,75	68,75	62,50
	Minkowski	81,25	71,88	65,63	59,38

Çizelge 4.3’de görüleceği üzere KNN sınıflandırıcıda kimi zaman eşit doğruluk oranı ile de karşılaşılabilir. Bu gibi durumlarda yalnızca doğruluk oranına değil devamında hesaplanan hassasiyet, özgüllük, kesinlik vb. değerlendirme ölçütleri ve kodun çalışma süresi de incelemeye dahil edilebilir. Örneğin; Min-Maks Normalizasyonda n=3 komşuluğunda Öklid, Cityblock ve Minkowski parametrelerinde KNN sınıflandırıcı eşit

doğrulukta çalışmıştır. $n=3$ komşuluğu için karmaşıklık matrisinden elde edilen hassasiyet, özgülük, kesinlik vb. değerlendirme ölçütleri de karşılaştırıldığında karmaşıklık matrislerinin aynı olması sebebi ile bu parametreleri de eşit değerde elde edilmiştir. Bu durumda hangisinin daha hızlı çalıştığı incelenebilir. Örneğin; $n=3$ komşuluğunda Öklid uzaklık bağıntısında 1.7631 saniyede sonuç elde edilirken, Cityblock uzaklık bağıntısında aynı sonuca 1.7032 saniyede ulaşılmaktadır, Minkowski uzaklık bağıntısında ise 1.7194 saniyede aynı sonuç elde edilmektedir. Dolayısıyla Min-Maks Normalizasyon ile KNN sınıflandırıcı yönteminde en yüksek doğrulukla ve en kısa sürede çalışan parametreler $n=3$ komşu sayısında, Cityblock uzaklık parametresidir denilebilir.

Naive Bayes sınıflandırma algoritmasında da KNN ve SVM'e benzer bir sonuç dikkat çekmektedir. Ham veri seti kullanılarak yapılan Naive Bayes sınıflandırıcıda Min-Maks normalizasyon sınıflandırma başarısını artırırken, Sıfır ortalama normalizasyon yöntemi sınıflandırıcı başarısını düşürmüştür. Naive Bayes sınıflandırıcıda normalizasyon uygulanmadığında sınıflandırıcı %71,88 doğruluk oranındayken, Min-Maks normalizasyon uygulandığında bu oran %78,13'e yükselmiştir. Sıfır ortalama normalizasyon ise %62,50 ile sınıflandırma performansını düşürmüştür.

Ham veriye öznitelik çıkarım fonksiyonları uygulanarak öznitelik sayısı 21'den 152'ye çıkarıldığında SVM sınıflandırıcıdan normalizasyon olmayan, Min-Maks normalizasyon yöntemi ile ve Sıfır ortalama normalizasyon yöntemi ile elde edilen sonuçlar Çizelge 4.4'te gösterilmiştir.

Çizelge 4.4. incelendiğinde Min-Maks normalizasyonun sınıflandırma başarısını önemli ölçüde arttırdığı gözlenmiştir. Sıfır ortalama normalizasyon yöntemi ise hiç normalizasyon yapılmayan duruma göre sınıflandırma başarısını düşürmüştür. Bu durum KNN ve Naive Bayes sınıflandırıcı sonuçları karşılaştırıldığında da aynı şekildedir. KNN sınıflandırıcıda $K=3, 4, 5$ ve 10 değerleri için Öklid, Cityblock, Chebychev ve Minkowski uzaklık bağıntıları için normalizasyon uygulanmadığında sınıflandırma sonucu en yüksek %96.88 iken, Min-Maks normalizasyon uygulandığında %100 doğruluk oranına ulaşılmıştır. Sıfır Ortalama normalizasyon yöntemi uygulandığında ise bu değer %90,63'e düşmüştür. Bu sonuçlardan yola çıkarsak bu veri setinde SVM, KNN ve Naive Bayes

sınıflandırma algoritmalarında Min-Maks normalizasyon, sınıflandırıcıların doğruluk oranını artırırken, Sıfır Ortalama normalizasyon yöntemi başarıyı düşürmektedir diyebiliriz.

Çizelge 4.4. 152 öznitelik ile SVM sınıflandırıcıdan elde edilen sonuçlar.

	Normalizasyon Yok	Min-Maks Normalizasyon	Sıfır Ortalama Normalizasyon
0=LİNEER	62,50	<u>96,88</u>	37,50
1=POLİNOMAL	62,50	62,50	62,50
2=RADYAL (RBF)	62,50	<u>96,88</u>	62,50
3=SIGMOİD	62,50	<u>96,88</u>	62,50
Ortalama doğruluk	62,50	88,29	56,25

4.2. Çapraz Doğrulama Yöntemlerinin Kıyaslanması

Sınıflandırma performansını etkileyen bir diğer husus da çapraz doğrulama yöntemidir. Bu çalışmada Birini Dışarda Bırakma Çapraz Doğrulama (Leave-one-out cross validation) ve KFold Çapraz Doğrulama (KFold Cross Validation) yöntemleri kullanılarak karşılaştırma yapılmış ve etkinlikleri kıyaslanmıştır.

Çapraz doğrulama yöntemlerinin uygulanma sebebi; algoritmanın aşırı öğrenmesini engellemek ve gerçekçi sonuçlar elde etmektir. Çapraz doğrulama yöntemleri ile verilerin bir kısmı eğitim seti olarak, kalan kısmı da eğitilen sistemin doğruluğunu test etmek için test set olarak kullanılmaktadır. Çapraz doğrulama işlemlerinin yapılmadığı durumlarda, hem eğitim hem de test set olarak aynı verilerin kullanılması gerekmekte, bu da sistemin ezbere düşmesine yani aşırı öğrenmesine ve gerçekçi olmayan yüksek doğruluk oranlarına ulaşılmasına neden olmaktadır.

Normalizasyon uygulanmadan, Min-Maks normalizasyon ve sıfır ortalama normalizasyon uygulanarak gerçekleştirilen denemelerde Min-Maks normalizasyonun KNN sınıflandırıcı başarısını olumlu yönde etkilediği tespit edilmiştir. Bu nedenle KNN sınıflandırıcıda ham verilerle yalnızca Min-Maks normalizasyon uygulanarak elde edilen farklı çapraz doğrulama yöntemlerinin sınıflandırma performansları Çizelge 4.5'te gösterilmiştir.

Çizelge 4.5. incelendiğinde Kfold ve Leaveout çapraz doğrulama yöntemleri sınıflandırıcı başarısını ayırt edici derecede etkilememiştir. Yalnızca Minkowski uzaklık bağıntısında $k=5$ komşuluk sayısında Kfold yöntemi, Leaveout yöntemine göre %9,37'lik başarı artışının sağlandığı görülmektedir. Bu durum ortalama bir doğruluk hesabı yapılacağı zaman Kfold'un Leaveout'a göre küçük bir miktar fazla doğrulukla çalıştığını gösterebilir. Yalnızca doğruluk oranından yola çıkılarak böyle bir yorum yapılması doğru olmayacaktır. Ancak hem Kfold hem de Leaveout çapraz doğrulama yöntemlerinden elde edilen karmaşıklık matrisleri de incelendiğinde neredeyse aynı matris sonuçlarının elde edildiği gözlenmiştir. Dolayısıyla, bu sonuçlardan hassasiyet, f-ölçütü ya da mcc değeri üzerinden de farklı bir yorum getirme imkanı bulunmamaktadır. Hangi yöntemin daha etkili olduğunun belirlenmesi ve diğer parametrelerdeki artışların incelenmesi için öznitelik çıkarma yöntemlerinden sonra elde edilen sınıflandırma başarılarının karşılaştırması daha net yorumlar yapabilmemize olanak sağlayacaktır.

Çizelge 4.5. KFold ve Leaveout çapraz doğrulama yöntemlerinin ham veri üzerinden KNN sınıflandırıcıda karşılaştırılması.

	KNN Uzaklık bağıntıları	KOMŞULUK SAYISI				
		3	5	7	9	
KFOLD ÇAPRAZ DOĞRULAMA	K=3	Öklid	<u>87,50</u>	78,13	75,00	75,00
		Cityblock	<u>87,50</u>	<u>87,50</u>	81,25	75,00
		Chebychev	81,25	81,25	75,00	68,75
		Minkowski	<u>87,50</u>	<u>87,50</u>	75,00	75,00
	K=4	Öklid	<u>87,50</u>	78,13	75,00	75,00
		Cityblock	<u>87,50</u>	<u>87,50</u>	81,25	75,00
		Chebychev	81,25	81,25	75,00	68,75
		Minkowski	<u>87,50</u>	<u>87,50</u>	75,00	75,00
	K=5	Öklid	<u>87,50</u>	78,13	75,00	75,00
		Cityblock	<u>87,50</u>	<u>87,50</u>	81,25	75,00
		Chebychev	81,25	81,25	75,00	68,75
		Minkowski	<u>87,50</u>	<u>87,50</u>	75,00	75,00
	K=10	Öklid	<u>87,50</u>	78,13	75,00	75,00
		Cityblock	<u>87,50</u>	<u>87,50</u>	81,25	75,00
		Chebychev	81,25	81,25	75,00	68,75
		Minkowski	<u>87,50</u>	<u>87,50</u>	75,00	75,00
LEAVEOUT ÇAPRAZ DOĞRULAMA	Öklid	<u>87,50</u>	78,13	75,00	75,00	
	Cityblock	<u>87,50</u>	<u>87,50</u>	81,25	75,00	
	Chebychev	81,25	81,25	75,00	68,75	
	Minkowski	<u>87,50</u>	78,13	75,00	75,00	

152 öznitelik ile sınıflandırma yapılarak çapraz doğrulama yöntemlerinin etkinliği tekrar kıyaslanmıştır. Çizelge 4.6. elde edilen bu sonuçları tek tek göstermektedir. Çizelge 4.5. ve Çizelge 4.6. karşılaştırıldığında öznitelik sayısının artmış olması sınıflandırma başarısını arttırmıştır. Ancak Kfold ve Leaveout çapraz doğrulama yöntemlerinin sınıflandırma başarısına olan etkilerini kıyaslayacak anlamlı ayrımlar gözlenememiştir. Kfold ve Leaveout çapraz doğrulama yöntemlerinden her ikisi de sınıflandırma başarısının artmasında eşit derecede etkilidir yorumu yapılabilir.

Çizelge 4.6. KNN sınıflandırıcıda Kfold ve Leaveout çapraz doğrulama yöntemlerinin 152 öznitelik kullanılarak karşılaştırılması.

		KNN Uzaklık bağıntıları	KOMŞULUK SAYISI			
			3	5	7	9
KFOLD ÇAPRAZ DOĞRULAMA	K=3	Öklid	96,88	90,63	90,63	87,50
		Cityblock	<u>100,00</u>	<u>100</u>	87,50	90,63
		Chebychev	50,00	37,50	37,50	37,50
		Minkowski	96,88	90,63	90,63	87,50
	K=4	Öklid	96,88	90,63	90,63	87,50
		Cityblock	<u>100,00</u>	<u>100</u>	87,50	90,63
		Chebychev	50,00	37,50	37,50	37,50
		Minkowski	96,88	90,63	90,63	87,50
	K=5	Öklid	96,88	90,63	90,63	87,50
		Cityblock	<u>100,00</u>	<u>100</u>	87,50	90,63
		Chebychev	50,00	37,50	37,50	37,50
		Minkowski	96,88	90,63	90,63	87,50
	K=10	Öklid	96,88	90,63	90,63	87,50
		Cityblock	<u>100,00</u>	<u>100</u>	87,50	90,63
		Chebychev	50,00	37,50	37,50	37,50
		Minkowski	96,88	90,63	90,63	87,50
LEAVEOUT ÇAPRAZ DOĞRULAMA	Öklid	96,88	90,63	90,63	87,50	
	Cityblock	<u>100,00</u>	<u>100</u>	87,50	90,63	
	Chebychev	50,00	37,50	37,50	37,50	
	Minkowski	96,88	90,63	90,63	87,50	

Naive Bayes sınıflandırma algoritmasında ise Leaveout ve Kfold çapraz doğrulama yöntemlerinin her ikisi de denenmiştir. Genel olarak aynı veri setleri üzerinde aynı normalizasyon yöntemleri uygulandığında Kfold, Leaveout'a göre daha başarılı sonuçlar elde edilmesini sağlamıştır. Sonuçların detayları Çizelge 4.7'de paylaşılmıştır.

Çizelge 4.7. Naive Bayes Sınıflandırma yönteminde 152 öznitelik ile yapılan sınıflandırma sonuçlarında çapraz geçерleme yöntemlerinin karşılaştırılması

Çapraz Doğrulama Yöntemleri	Normalizasyon Yöntemleri		
	Normalizasyon Yok	Min-Maks Norm.	Sfır Ortalama Norm.
Leaveout	78,13	71,88	68,75
Kfold K=3	75,00	75,00	75,00
Kfold K=4	78,13	75,00	68,75
Kfold K=5	<u>84,38</u>	78,13	75,00
Kfold K=10	81,25	75,00	71,88

4.3. Sınıflandırma Yöntemlerinin Kıyaslanması

Ham veri seti için normalizasyon kombinasyonları içinde en başarılı sınıflandırma yöntemi KNN'dir. SVM ve Naive Bayes sınıflandırıcıların sınıflandırıcı doğrulukları neredeyse eşitken, SVM'in normalizasyon yapılmadığı durumda %3,12 fazla doğru tahmini vardır.

Çizelge 4.8'de SVM, KNN ve Naive Bayes sınıflandırma yöntemlerinden 152 öznitelik ile elde edilen en yüksek doğruluk oranları verilmiştir. Sınıflandırma algoritmalarının üçünde de leaveout çapraz doğrulama yöntemi kullanılmıştır.

Çizelge 4.8. Leaveout çapraz doğrulama ile 152 öznitelik kullanılarak elde edilen sınıflandırma en yüksek doğruluk oranları.

Sınıflandırma Yöntemi	Uygulanan Normalizasyon Yöntemi		
	Normalizasyon Yok	Min-Maks Normalizasyon	Sfır Ortalama Normalizasyon
SVM	62,50	96,88	62,50
KNN	96,88	100	90,63
Naive Bayes	78,13	71,88	68,75

Sonuç olarak aynı normalizasyon yöntemi ve aynı çapraz doğrulama yöntemi kullanıldığında aynı veri seti üzerinden elde edilen sınıflandırıcı sonuçları farklılık göstermiştir. Normalizasyon seçeneklerinin tamamında en yüksek doğruluk oranına KNN

sınıflandırıcı ile ulaşılmıştır. SVM sınıflandırıcıda Min-Maks normalizasyon yöntemi sınıflandırma başarısını anlamlı ölçüde artırırken, Naive Bayes sınıflandırıcıda küçük bir miktar düşürmüştür. Sınıflandırma yöntemlerinin tümünde sıfır ortalama normalizasyon uygulandığında sınıflandırma başarısının düştüğü gözlenmiştir. Naive Bayes sınıflandırma algoritması özelinde sonuçlar incelendiğinde normalizasyon yöntemlerinin her ikisi de başarıyı olumsuz etkilemiştir.

Çizelge 4.8 yalnızca doğruluk oranları ile oluşturulmuştur. Fakat sınıflandırıcı başarısını belirleyen tek parametre doğruluk değeri değildir. Diğer değerlendirme ölçütlerinin de her bir sınıflandırıcı için incelenmesi, karşılaştırılması ve yorumlanması gerekmektedir.

Çizelge 4.9. Min-Maks normalizasyon ve Leaveout çapraz doğrulama ile 152 öznitelik kullanılarak elde edilen en yüksek sınıflandırma performansları.

	Uygulanan Normalizasyon Yöntemi		
	Min-Maks Normalizasyon		
	Doğruluk	F-ölçütü	mcc
SVM	96,88	0,96	0,9608
KNN	100	1	1
Naive Bayes	71,88	0,7273	0,5134

Çizelge 4.9’da 152 öznitelik ile üç sınıflandırma algoritması için elde edilen sonuçların değerlendirme ölçütleri bulunmaktadır. KNN sınıflandırıcıda maksimum doğruluğun elde edildiği kombinasyonda karmaşıklık matrisi [12 0; 0 20] olarak elde edilmiştir. Yani algoritma Parkinson hastalarının tamamını hasta tahmin etmiş, sağlıklıların da tamamını sağlıklı olarak tahmin etmiştir.

Mcc değerinin sistem başarısı arttıkça +1 değerine yaklaşması beklenir. Karmaşıklık matrisleri ve doğruluk oranları incelendiğinde en iyi performansı KNN sınıflandırıcı vermiştir. Bu durumu F-ölçütü ve mcc değerinin 1 elde edilmesi de desteklemektedir. KNN’den sonraki en iyi sınıflandırıcı SVM’dir. SVM sınıflandırıcının 0,9608’lik mcc değeri ve 0,96’lık F-ölçütü değeri yüksek değerler olup sınıflandırma performansının iyi olduğunu göstermektedir.

F-ölçütünde de mcc değerinde olduğu gibi sistem başarı performansı arttıkça +1 değerine yakın bir değer elde edilmesi beklenir. KNN sınıflandırıcıda F-ölçütünün 1 elde edilmesi sınıflandırıcının yüksek doğruluk ile çalıştığını göstermektedir.

SVM sınıflandırıcının en yüksek doğruluğu Min-Maks normalizasyon yöntemi ve Leaveout çapraz doğrulama yöntemi ile gösterdiği tespit edilmiştir. SVM sınıflandırıcıda 0=linear, 2=RBF, 3=sigmoid yaklaşımlarının üçünde aynı sonuç elde edilmiş ancak bu üçünden en hızlı çalışan parametrenin 0,212675 saniye ile 3=sigmoid yaklaşımı olduğu sonucu elde edilmiştir.

KNN sınıflandırıcının en yüksek doğruluğu Min-Maks normalizasyon yöntemi ve Kfold ve Leaveout çapraz doğrulama yönteminin ikisi ile de gösterdiği tespit edilmiştir. Burada KNN sınıflandırıcının Cityblock uzaklık bağıntısı ile ve 3 ve 5 komşu sayısı ile en yüksek doğruluk değeri elde edilmiştir. Kfold çapraz doğrulamada K=3, 4, 5, 10 değerlerinin tamamı için sınıflandırıcı çalıştırılmış ve tüm K değerleri için aynı uzaklık bağıntılarında ve aynı komşuluk değerlerinde eşit sonuçlar elde edilmiştir. Tüm kombinasyonlar incelendiğinde aynı en yüksek sonuçlar içinde en hızlı çalışan kombinasyon KNN sınıflandırıcı için: 0,896976 saniye ile Min-Maks Normalizasyon, Kfold'da K=3 değeri için, 5 komşu sayısı ve Cityblock uzaklık parametresinin kullanıldığı kombinasyondur.

Naive Bayes Sınıflandırıcının en yüksek doğruluğu normalizasyon yapılmadan, Kfold çapraz geçiş yöntemi ile elde edilmiştir. Kfold'da K=5 değerinde en iyi performans gözlenmiştir. Bu sınıflandırma sonucuna 6.557659 saniyede ulaşılmıştır. Bu süre SVM ve KNN sınıflandırıcıların sınıflandırma sürelerinin çok üzerinde bir değerdir.

4.4. Öznitelik Seçme Algoritmalarının Uygulanması

Öznitelik sayısı, sınıflandırma süresini ve sınıflandırma performansını doğrudan etkileyen önemli faktörlerdendir. Öznitelik çıkarım fonksiyonları kullanılarak 21 olan öznitelik sayısı 152'ye çıkarılmıştır. Çalışmanın bu aşamaya kadar açıklanan bölümlerinde hem 21 öznitelik ile hem de 152 öznitelik ile elde edilen sonuçlar paylaşıp karşılaştırılmaya çalışılmıştır.

Çizelge 4.10. 21 öznitelik ve 152 öznitelik ile elde edilen sınıflandırma performanslarının karşılaştırılması.

Sınıflandırma Yöntemi	Öznitelik Sayıları İle Elde Edilen Doğruluk Oranları	
	21 Adet Öznitelik	152 adet öznitelik
SVM	78,13	96,88
KNN	87,50	100
Naive Bayes	78,13	84,38

Çizelge 4.10'de verilmiş olan sınıflandırma performansları karşılaştırıldığında öznitelik sayısının arttırılması sınıflandırıcıların doğruluk oranlarının artmasını da sağlamıştır. Hatta yalnızca öznitelik sayısı arttırılarak KNN sınıflandırıcıda %100 başarı oranına ulaşılmıştır. Ancak bu tabloda 152 öznitelik kullanılarak yapılan sınıflandırma sonuçlarının güvenilirliğinin test edilmesi, daha az sayıda öznitelik ile beklenen performans sonuçlarına ulaşılabilmesi ve %100'e ulaşmayan sınıflandırıcı doğruluklarının daha da arttırılabilmesi için öznitelik seçim yöntemlerinin uygulanması faydalı olacaktır.

Öznitelik seçme yöntemleri kullanılarak belirleyici olan özniteliklerin tespiti yapılabilmektedir. Yani çıkarılan 152 öznitelik arasından en belirleyici olanların seçilerek sınıflandırıcılardaki başarılarının test edilmesi denenmiştir. 152 özniteliğin tamamının kullanılması yerine sadece az sayıda belirleyici özniteliklerin kullanılması hem işlem sayısının azalmasını sağlar hem de sınıflandırıcıya olumsuz katkı sağlayan özniteliklerin elenerek başarının arttırılmasını sağlar. Böylelikle daha az sayıda öznitelik kullanılarak daha fazla başarı elde etmek hedeflenmiştir. Bu durum hem sınıflandırma performansının arttırılması hem de sınıflandırıcıların harcadıkları süre ve enerjiyi minimum seviyeye düşürmeye katkı sağlar. Yani daha kısa sürede, daha az öznitelik ile daha yüksek başarı elde edilebilir.

Bu amaç doğrultusunda bu tez çalışmasında üç farklı öznitelik seçme yöntemi uygulanmıştır. Hasta-sağlıklı ayrımının yapılmasında en belirleyici olan öznitelikler tespit edilerek aralarından en etkili olan ilk 15 öznitelik kaydedilmiş ve bu 15'er öznitelik üç farklı sınıflandırma algoritmasında çalıştırılmıştır.

Çizelge 4.11. 152 öznitelik arasından belirleyici olan ilk 15 özniteliğin farklı yöntemlerle belirlenmesi.

	Normalizasyon Olmadan Öznitelik Seçme Yöntemleri Sonuçları			Min-Maks Norm. İle Öznitelik Seçme Yöntemleri Sonuçları		
	LASSO	mRmR	RELIEF	LASSO	mRmR	RELIEF
1	29	45	48	9	15	59
2	43	48	72	34	45	42
3	44	75	43	42	75	109
4	46	105	42	48	105	72
5	71	135	152	59	135	107
6	78	138	109	76	138	76
7	79	139	36	102	139	117
8	82	43	89	107	147	48
9	89	44	35	109	59	20
10	90	46	102	111	109	50
11	96	125	50	117	42	5
12	103	140	96	-	72	9
13	109	62	132	-	117	4
14	129	129	44	-	76	77
15	140	86	41	-	20	6

LASSO ile gerçekleştirilen öznitelik seçiminde yalnızca 11 öznitelik seçilebilmiştir.

Öznitelik seçiminde tercih edilen normalizasyon yöntemi de etkili olmuştur. Çizelge 4.11’de normalizasyon olmadan seçilen ilk 15 özniteliğin sıra numarası ve Min-Maks normalizasyon ile seçilen ilk 15 özniteliğin sıra numarası verilmiştir. Normalizasyon yönteminin varlığı ya da uygulanan farklı normalizasyon yöntemleri farklı özniteliklerin seçilmesine de neden olmaktadır.

Özniteliklerin sıra numaralarında 1-31 arası yürüyüşün kinetik verileri, 32-61 arası yürüyüşte kinematik kalça açısı, 62-91 arası yürüyüşte kinematik diz açısı, 92-121 arası yürüyüşte kinematik ayak bileği açısı, 122-151 arası yürüyüşün temel verileri, 152 ise hastanın vücut kitle indeksi verisidir. Çizelge 4.11’de verilen öznitelik numaraları dahil oldukları gruba göre renklendirilmiştir. Kinetik veriler pembe, kinematik veriler sarı, temel yürüyüş verileri yeşil ve VKİ verisi de turuncu renk ile işaretlenmiştir. Bu durumda normalizasyon olmayan durumda LASSO ile 1 kinetik, 12 kinematik, 2 temel yürüyüş verisi seçilmiştir. mRmR ile 9 kinematik, 6 temel yürüyüş verisi seçilmiştir. RELIEF ile de 13 kinematik, 1 temel yürüyüş verisi, 1 de VKİ verisi seçilmiştir. Yani normalizasyon yöntemi uygulanmayan verilerde kinematik verilerin sınıflandırmada daha belirleyici olduğu tespit edilmiştir. Min-Maks normalizasyon yöntemi uygulanan veriler arasından seçim yapıldığında LASSO yöntemi ile 1 kinetik, 10 kinematik veri, mRmR yöntemi ile 2 kinetik, 9 kinematik ve 4 temel yürüyüş verisi seçilmiştir. RELIEF ile 5 kinetik, 10 kinematik veri seçilmiştir. Min-Maks normalizasyon yöntemi uygulanan veriler arasından yapılan seçim işleminde de kinematik verilerin ağırlığı dikkat çekmektedir. Yani her iki durumda da kinematik veriler en belirleyici veriler arasındadır. Dolayısıyla yürüyüş ile Parkinson hastalığının tespitinde özellikle kinematik verilerin tercih edilmesi tahmin doğruluğunu arttıracaktır.

Kinematik veriler yürüyüşte gövde, pelvis, bacak ve ayaklardaki eklemelerde oluşan açılar, hızlar ve ivmelerin sayısal verileridir. Hastalığın dopamin eksikliğinden kaynaklı olarak hareket mekanizmasını olumsuz etkilediği bilinmektedir. Dopamin eksikliğinden kaynaklı olarak eklemlerle sağlıklı kişilerdeki açılarda açılmaz, hareket yavaşlar, ani ivmelenmeler görülür. Klinikte gözlenen bu bulgular öznitelik seçiminde tespit edilen özniteliklerle örtüşmektedir.

Bu çalışmada öznitelik seçiminden önce elde edilen sınıflandırma sonuçları değerlendirildiğinde, SVM ve KNN sınıflandırıcılar için Min-Maks normalizasyon uygulandığında daha yüksek doğruluk elde edildiğinden, öznitelik seçimi sonucunda elde edilen ilk 15 öznitelik için yalnızca Min-Maks Normalizasyon sonucunda elde edilen özniteliklerin sınıflandırma sonucuna etkisi değerlendirilmiştir.

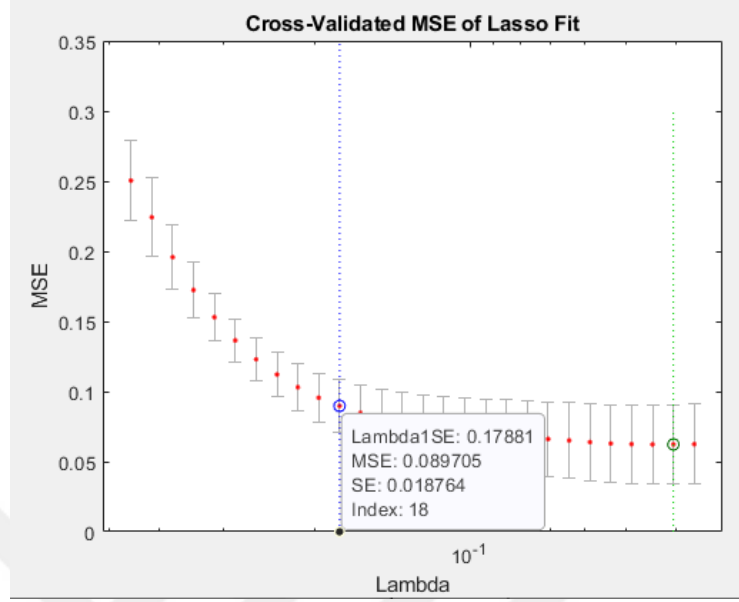
Naive Bayes sınıflandırıcıda ise normalizasyonun sınıflandırma performansını düşürdüğü sonucuna varıldığından, Naive Bayes sınıflandırıcı için hem normalizasyon yapılmadan hem de Min-Maks normalizasyon ile seçilen özniteliklerin sınıflandırma performansları irdelenmiştir.

4.4.1. LASSO (Least Absolute Shrinkage and Selection Operator- En Küçük Mutlak Daralma ve Seçme Operatörü) Yöntemi İle Öznitelik Seçimi

Yüksek belirleyiciliğe sahip özniteliklerin doğru, yanıltıcı özelliğe sahip özniteliklerin yanlış olarak seçildiği bir yöntemdir. Böylelikle en belirleyici doğru öznitelikler tespit edilerek, bu öznitelikler kullanıldığında sınıflandırma performansının artırılması amaçlanır.

152 öznitelik içinden LASSO yöntemi ile seçilen ilk 15 öznitelik Çizelge 4.11’de verilmiştir. Min-Maks normalizasyon yapılarak oluşturulan seçim işlemi yalnızca 11 öznitelik seçilebilmiştir. Normalizasyon uygulanmadan yapılan seçimlerde kinetik verilerin fast approximate entropi değeri (29), kalça açısının çarpıklığı (43), kalça açısının basıklığı (44), kalça açısının Higuchi değeri (46), diz açısının aralık değeri (71) en belirleyici öznitelikler olarak seçilmiştir. Min-Maks normalizasyon uygulandığında ise; kinetik verilerin orta (medyan) değeri (9), maximum fleksiyonda kalça açısı değeri (34), kalça açısının entropi değeri (42), kalça açısının Sample Entropi değeri (48) ve kalça açısının fast approximate Entropi değeri (59) en belirleyici öznitelikler olarak seçilmiştir.

Şekil 4.1’de Min-Maks normalizasyon ile çapraz doğrulamalı LASSO seçiminde MSE grafiği ve değeri verilmiştir. LASSO, doğrusal regresyona dayalı bir yöntem olduğundan hesaplanan ortalama kare hata (MSE) değeri anlamlıdır. MSE değeri 0,089705 olarak hesaplanmıştır. MSE değerleri içinde sıfıra en yakın olan değer en etkin olan model olarak yorumlanır. LASSO’da farklı lambda değerleri ile öznitelik alt kümeleri oluşturulmuş 100 farklı model içinden 18.si sıfıra en yakın olan MSE değerine sahiptir.



Şekil 4.1. Çapraz doğrulamalı Min-Maks normalizasyon ile LASSO seçiminde MSE değeri ve grafiği.

4.4.2. mRmR (Minimum Artıklık Maksimum İlişki) İle Öznitelik Seçimi

mRmR, çalışma hızı ve başarı oranı yüksek olan entropi tabanlı bir öznitelik seçme yöntemidir. Entropi tabanlı bir yöntem özniteliklerin filtrelenerek seçilmesini sağlar.

mRmR yöntemi ile seçilen ilk 15 öznitelik Çizelge 4.11’de verilmiştir. mRmR yöntemi ile normalizasyon uygulanmadan yapılan seçimlerde kalça açısının Kolmogorov değeri (45), Sample Entropi değeri (48), diz açısının Kolmogorov değeri (75), ayak bileği açısının Kolmogorov değeri (105) ve temel verilerin Kolmogorov değeri (135) en belirleyici öznitelikler olarak seçilmiştir. Min-Maks normalizasyon uygulandığında ise; kinetik verilerin Kolmogorov değeri (15), kalça açısının Kolmogorov değeri (45), diz açısının Kolmogorov değeri (75), ayak bileği açısının Kolmogorov değeri (105) ve temel verilerin Kolmogorov değeri (135) en belirleyici öznitelikler olarak seçilmiştir.

mRmR yöntemi ile seçilen öznitelikler incelendiğinde verilerden Kolmogorov fonksiyonu ile elde edilen yeni verilerin sıklıkla seçildiği dikkat çekmektedir. mRmR öznitelik seçiminde Kolmogorov değerleri belirleyici olmuştur denilebilir.

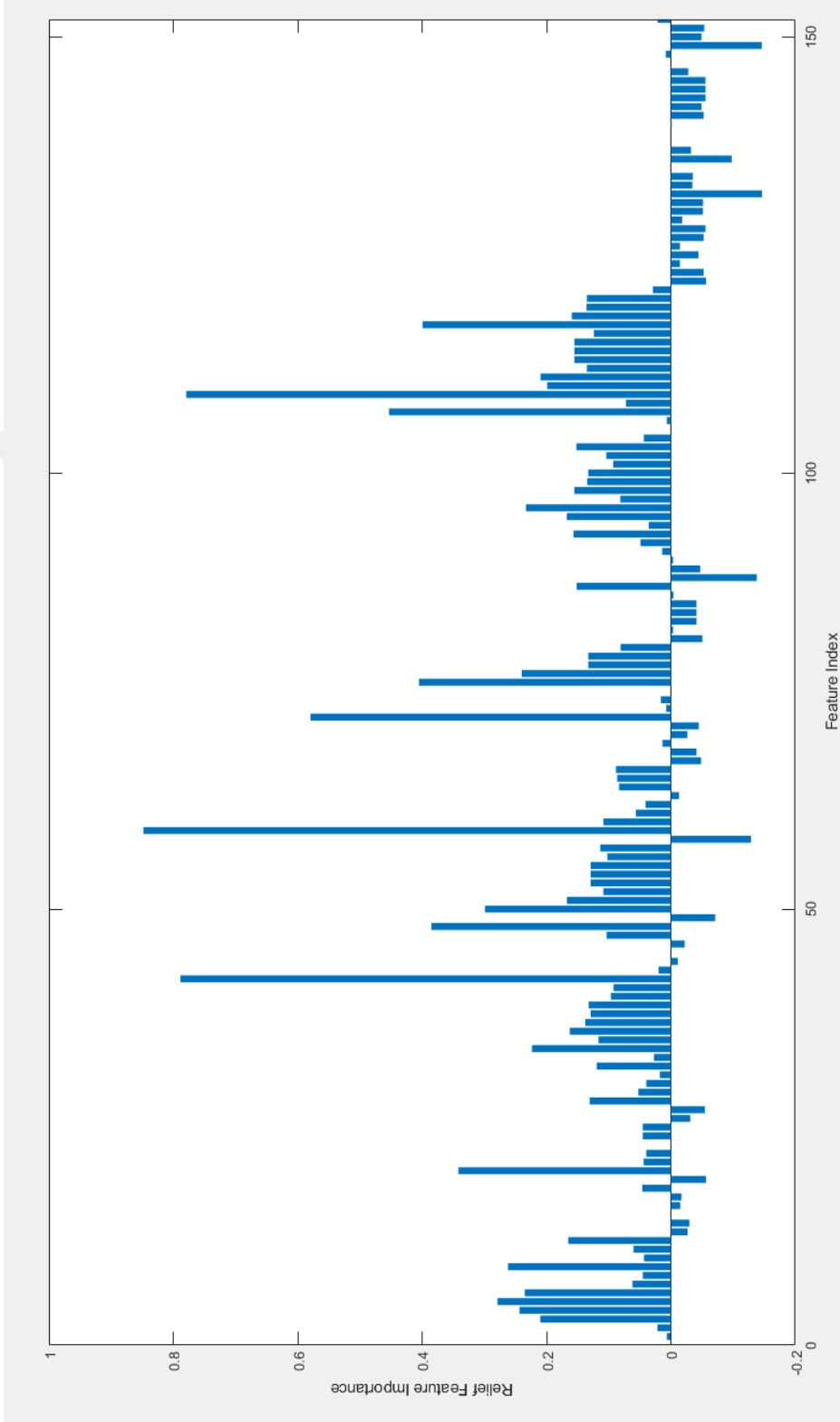
4.4.3. RELIEF İle Öznitelik Seçimi

KNN sınıflandırma algoritmasının çalışma prensibine benzer bir özellik gösteren RELIEF yönteminde, belirleyici özelliklere sahip özniteliklere Öklid uzaklık ölçümü ile hesaplanan değerler, özniteliklere ağırlık olarak atanır. Bu çalışmada elde edilen 152 özniteliğin RELIEF öznitelik seçim yöntemi ile seçilmesinde oluşturulan ağırlıklandırma grafiği Şekil 4.2’de gösterilmiştir.

RELIEF yöntemi ile öznitelik seçiminde en düşük performans gösteren ya da bir başka deyişle sınıflandırıcıya olumsuz katkı sağladığı düşünülen özniteliklere sıfır değeri atanarak o öznitelikler veri kümesinden çıkartılır.

RELIEF yöntemi ile seçilen ilk 15 öznitelik Çizelge 4.11’de verilmiştir. RELIEF yöntemi ile normalizasyon uygulanmadan yapılan seçimlerde kalça açısının Sample Entropi değeri (48), diz açısının Entropi değeri (72), kalça açısının çarpıklığı (43), kalça açısının entropi değeri (42) ve vücut kitle indeksi (152) en belirleyici öznitelikler olarak seçilmiştir. Min-Maks normalizasyon uygulandığında ise; kalça açısının fast approximate Entropi değeri (59), kalça açısının entropi değeri (42), ayak bileği açısının multiscale sample entropi değeri (109), diz açısının Entropi değeri (72), Ayak bileğinin Katz fraktal boyut değeri (107) en belirleyici öznitelikler olarak seçilmiştir.

RELIEF yöntemi hem normalizasyon uygulanmadan yapılan seçimde hem de Min-Maks normalizasyon uygulanmış seçimde 42 ve 72 numaralı verileri ortak olarak seçmiştir. Bunlar: kalça açısının entropi değeri (42) ve diz açısının Entropi değeri (72)’dir.



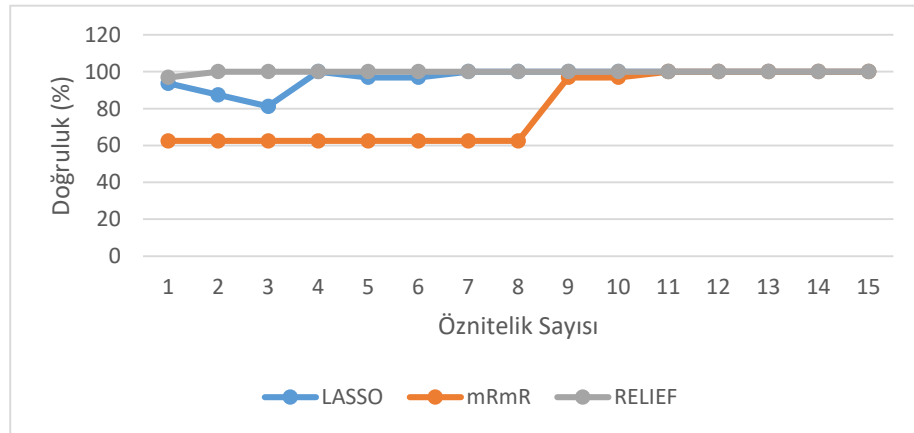
Şekil 4.2. RELIEF öznelik seçim yöntemi ile 152 öznelik ağırlıklandırılmasına dair grafik

4.5. Öznitelik Seçme Algoritmalarının Sınıflandırma Sonuçlarına Katkılarının Karşılaştırılması

LASSO, mRmR ve RELIEF yöntemleri ile seçilen 15 öznitelik SVM, KNN ve Naive Bayes sınıflandırıcılarda ayrı ayrı çalıştırılmıştır.

Şekil 4.3'te verilen grafikte LASSO, mRmR ve RELIEF öznitelik seçme yöntemleri ile seçilen 15 özniteliğin SVM sınıflandırıcı algoritmasında çalıştırılması sonucunda elde edilen doğruluk oranları yer almaktadır. Grafığe öznitelik sayısı birer birer arttırılarak 1, 2, 3, ... , 15 adet özniteliğin ayrı ayrı çalıştırılması ile ulaşılmıştır. Ayrıca SVM sınıflandırıcı yönteminde öznitelik sayısı arttıkça sınıflandırıcının çalışma süresinin de arttığı görülmüştür. Bu grafikte elde edilen sonuçların tümü verilere Min-Maks normalizasyon ve Leaveout çapraz doğrulama yöntemi uygulanarak elde edilmiştir.

152 öznitelik, Min-Maks normalizasyon ve Leaveout çapraz doğrulama yöntemi ile SVM sınıflandırıcıda elde edilen en yüksek doğruluk değeri %96,88, sınıflandırıcının çalışma süresi ise bu doğruluk oranında en az 0,212675 saniyedir. Farklı yöntemlerle seçilen 15 öznitelik kullanılarak yapılan sınıflandırma sonucu %100 gibi en yüksek orana ulaşmıştır. Süre de 0,133137 saniyeye düşmüştür.



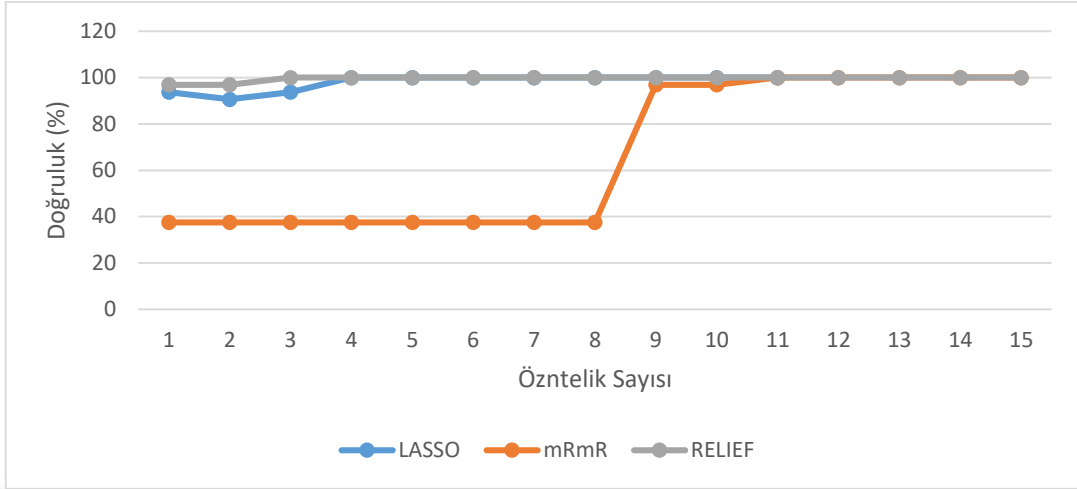
Şekil 4.3. SVM sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları.

Sonuç olarak SVM sınıflandırıcı için LASSO, mRmR ve RELIEF yöntemleri ile seçilen 15'er öznitelikle yapılan sınıflandırma yüksek performans göstermiştir. Şekil 4.3'te de görüldüğü üzere RELIEF ile seçilen 3 öznitelikten itibaren sınıflandırıcı

maksimum doğruluğa ulaşmıştır. Dolayısıyla SVM sınıflandırıcı için RELIEF seçim yöntemi en başarılı performansı göstermiştir denilebilir.

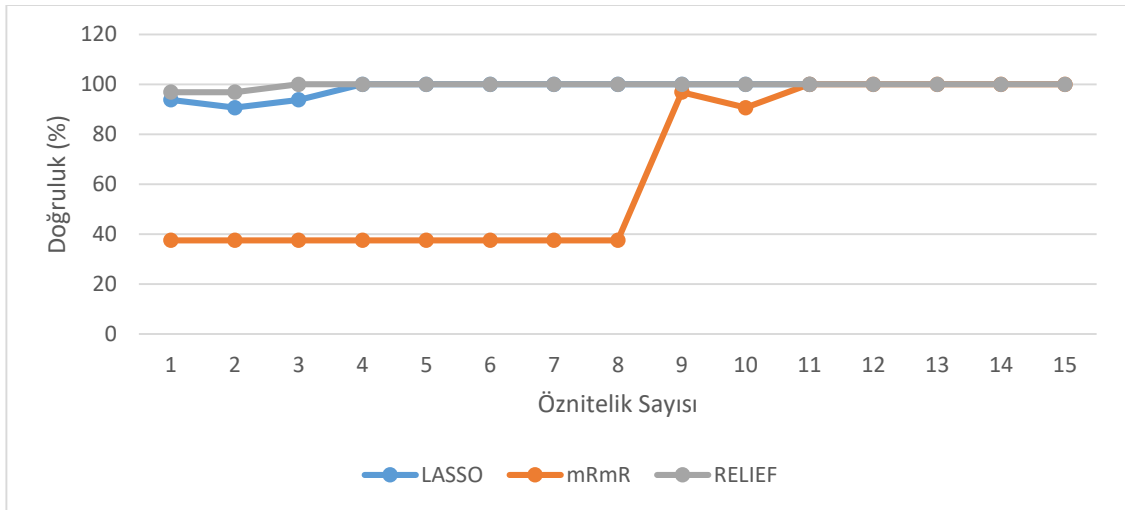
KNN sınıflandırıcı için de üç öznitelik seçme algoritmasından gelen 15 öznitelik çalıştırılarak sonuçlar grafiklendirilmiştir. KNN için hem Leaveout hem de Kfold çapraz doğrulama yöntemleri ayrı ayrı çalıştırılmıştır. KNN sınıflandırıcıda en yüksek doğruluk oranı Min-Maks normalizasyon yöntemi ile elde edilmiştir. Bu sebeple seçilmiş özniteliklerle yapılan sınıflandırmada da Min-Maks normalizasyon yöntemi uygulanmıştır.

Şekil 4.4'te verilen grafik KNN sınıflandırıcıda seçilen 15'er özniteliğin Leaveout çapraz doğrulama yöntemi ile sınıflandırma doğruluk değerleri bulunmaktadır. 152 öznitelik ile elde edilen %100'lük doğruluk oranına, RELIEF yöntemi ile seçilen 3 öznitelik ile ulaşılmıştır. Bu durum daha az öznitelik kullanılarak daha hızlı ve performansı yüksek sınıflandırma işlemi yapıldığını göstermektedir. 152 öznitelik ile yapılan sınıflandırma işleminde %100'lük doğruluk oranına min-maks normalizasyon, Kfold ve Leaveout çapraz doğrulama yöntemlerinin her ikisi ile de ve KNN sınıflandırıcının 3 ve 5 komşuluk değerleri, Cityblock uzaklık bağıntısı ile 0,896976 saniyede ulaşılmıştır. RELIEF seçim yöntemi ile elde edilen 15 öznitelikle yapılan sınıflandırmada da en yüksek doğruluk oranına KNN sınıflandırıcının 3 ve 5 komşuluk değerleri, Cityblock uzaklık bağıntısı ile 0,415394 saniyede ulaşılmıştır. Daha az sayıda öznitelik kullanımı doğal olarak burada da sınıflandırma süresini neredeyse yarı yarıya kısaltmıştır.



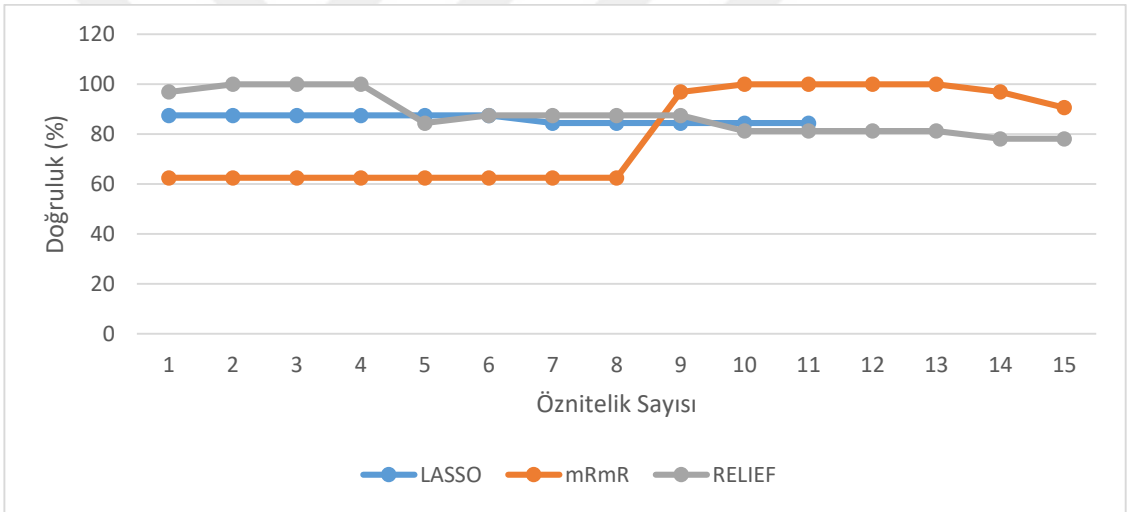
Şekil 4.4. KNN sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Leaveout çapraz doğrulama ile).

Şekil 4.5.'te KNN sınıflandırıcı da Min-Maks normalizasyon ve Kfold (K=10) ile 3 farklı yöntemle seçilen 15'er özniteliğin doğruluk oranlarının karşılaştırması bulunmaktadır. Şekil 4.4'te olduğu gibi burada da en yüksek doğruluk oranına en hızlı RELIEF algoritması ile seçilen öznitelikler kullanılarak ulaşılmıştır. Sınıflandırma süresi bu yöntemde daha da kısalmış, 0,253806 saniyede sınıflandırma işlemi sonlanmıştır.



Şekil 4.5. KNN sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Kfold (K=10) çapraz doğrulama ile).

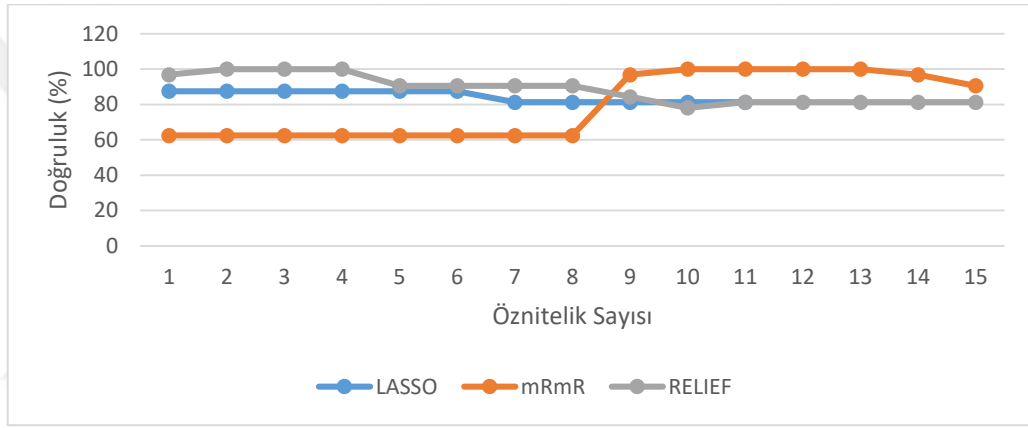
Son olarak Naive Bayes sınıflandırıcıda da aynı işlemler yapılmıştır. Şekil 4.6’da Naive Bayes sınıflandırıcıda ilk 15’er öznitelikliğin sınıflandırma doğruluklarının Leaveout çapraz doğrulama yöntemi ile elde edilen sonuçları bulunmaktadır. Buradaki sonuçlarda Min-Maks normalizasyon yapılmış verilerle Leaveout çapraz doğrulama yönteminde Naive Bayes sınıflandırıcı doğruluk oranları görülmektedir. RELIEF öznitelik seçme yöntemi önceki sınıflandırıcılarda olduğu gibi en az öznitelik ile en yüksek doğruluk oranına (%100) ulaşılmasını sağlamıştır. Ancak RELIEF’in seçtiği devamındaki öznitelikler sınıflandırma başarısını %78,13’e kadar düşürmüştür. Bu değer 152 öznitelik ile yapılan Naive Bayes sınıflandırmada Min-Maks normalizasyon ve Leaveout çapraz doğrulama ile elde edilen maksimum doğruluğa eşittir.



Şekil 4.6. Naive Bayes sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Leaveout çapraz doğrulama ile).

Naive Bayes sınıflandırıcıda Kfold çapraz doğrulama ile yapılan sınıflandırma sonuçları da Şekil 4.7’de görülmektedir. Naive Bayes sınıflandırma yönteminde 152 öznitelik ile yapılan önceki sınıflandırma sonuçlarında en yüksek doğruluk değerine Min-Maks Normalizasyonda Kfold çapraz doğrulama yöntemi ile K=5 değeri için %78,13 olarak elde edilmişti. LASSO, mRmR ve RELIEF öznitelik seçme yöntemleri ile seçilen 15’er öznitelik Naive Bayes sınıflandırıcıda çalıştırıldığında %100 doğruluk değerlerine ulaşılabilmiştir. LASSO’da en yüksek doğruluk değeri %87,50 olarak elde edilmiştir. mRmR ve RELIEF yöntemlerinde seçilen öznitelikler Naive Bayes sınıflandırıcıda çalıştırıldığında ise doğruluk değeri %100’ü görmüştür. Ancak RELIEF ile seçilen

özniteliklerden ilk ikisi, ilk üçü ve ilk dördü maksimum doğruluk değerine ulaşılmasını sağlarken ilk 5 öznitelikten itibaren yapılan sınıflandırma doğruluğu düşme eğilimi göstermiştir. Ancak buna rağmen doğruluk değeri RELIEF’te seçilen 15 öznitelikle yapılan sınıflandırma oranı %78,13’ün altına düşmemiştir. En anlamlı grafik burada mRmR ile seçilen özniteliklerden elde edilmiştir. Doğruluk oranı artış göstermiş bir süre sabit devam etmiş ilk 13 öznitelikten sonra düşme eğilimi göstermiştir. Bu durumda 14. Öznitelikten itibaren seçilen özniteliklerin arasında sınıflandırma başarısını düşüren veriler bulunmaktadır denilebilir.



Şekil 4.7. Naive Bayes sınıflandırıcıda üç yöntemle seçilen 15 öznitelikten elde edilen doğruluk sonuçları (Kfold (K=5) çapraz doğrulama ile).

4.6. Öznitelik Seçme Algoritmalarından Elde Edilen Ortak Öznitelikler ve Sınıflandırıcılara Etkileri

LASSO, mRmR ve RELIEF öznitelik seçme yöntemleri ile belirleyici olan 15’er öznitelik seçilmiştir. Her bir yöntemin öznitelik seçme yaklaşımı farklı olduğundan seçtikleri öznitelikler de farklıdır. Öyleyse bu öznitelikler içinden ortak olanlar kullanılarak yapılacak sınıflandırma sonuçları da değerlendirilmelidir.

Öznitelik seçme yöntemlerinde seçim esnasında normalizasyon uygulanıp uygulanmamasına göre de sonuçlar fark etmektedir. Aşağıda herhangi bir normalizasyon yöntemi uygulanmadan seçilen ve Min-Maks normalizasyon uygulanarak seçilen özniteliklerle yapılan öznitelik seçimleri, üç yöntemle seçilen farklı özniteliklerden ortak

olanların belirlenmesi ve bu ortak özniteliklerin de sınıflandırma performansına etkileri ayrıca değerlendirilmiştir.

4.6.1. Normalizasyon yapılmadan seçilen ortak öznitelikler ve sınıflandırıcılardaki etkinlikleri

Çizelge 4.12’de Normalizasyon yapılmadan seçilen ortak öznitelikler ve bu öznitelikleri seçen seçme algoritmaları verilmiştir. Çizelgeye göre 1 ve 2. Sırada bulunan 43 ve 44. öznitelikler LASSO, mRmR ve RELIEF seçim yöntemlerinden üçünde de seçilmiştir. 3-9. Sıra arasındaki öznitelikler de iki seçme yöntemi tarafından seçilen özniteliklerdir. Bu sonuçlar yorumlandığında üç yöntemde de seçilen özniteliklerin kinematik veriler olduğu dikkat çekmektedir. İki yöntemde seçilen verilerde de yine kinematik verilen çoğunluğu dikkat çekmektedir, 7’de 5’i kinematik parametrelerdir.

Çizelge 4.12. Normalizasyon uygulanmadan yapılan seçim işlemlerindeki ortak öznitelikler ve seçen yöntemlerin tablo şeklinde gösterimi.

	Öznitelik	Özniteliği seçen yöntem
1	Kalça açısının çarpıklığı (43)	LASSO, mRmR ve RELIEF
2	Kalça açısının basıklığı (44)	LASSO, mRmR ve RELIEF
3	Kalça açısının Higuchi değeri (46)	LASSO ve mRmR
4	Kalça açısının sample entropi değeri (48)	MRmR ve RELIEF
5	Diz açısının fast approximate entropi değeri (89)	LASSO ve RELIEF
6	Ayak bileği açısının minimum değeri (96)	LASSO ve RELIEF
7	Ayak bileği açısının Multiscale sample entropi değeri (109)	LASSO ve RELIEF
8	Temel yürüyüş verilerinin orta (medyan) değeri (129)	LASSO ve mRmR
9	Temel yürüyüş verilerinin harmonik ortalaması (140)	LASSO ve mRmR

Ortak olarak seçilen bu 9 öznitelik farklı kombinasyonlarla sınıflandırıcılarda çalıştırılmış ve sonuçları karşılaştırılmıştır. Bu kombinasyonlar:

- Grup 1: Yalnızca 43 ve 44. Özniteliklerin kullanılarak sınıflandırma,
- Grup 2: 46, 48, 89, 96, 109, 129 ve 140. Öznitelikler kullanılarak sınıflandırma,
- Grup 3: 43, 44, 46, 48, 89, 96, 109, 129 ve 140. Öznitelikler kullanılarak sınıflandırma şeklindedir.

SVM, KNN ve Naive Bayes sınıflandırıcılarda Grup 1, Grup 2 ve Grup 3 ile elde edilen sonuçlar aşağıda tablo şeklinde verilmiştir.

Çizelge 4.13. Normalizasyon yapılmadan seçilen ortak öznitelikler ve farklı sınıflandırıcılardaki performansları.

	Grup 1	Grup 2	Grup 3
SVM	78,13	65,63	65,63
KNN	87,50	84,38	90,63
Naive Bayes	87,50	78,13	84,38

Çizelge 4.13'te SVM ve KNN'in tüm sınıflandırma parametrelerinden elde edilen en yüksek doğruluk oranları verilmiştir. Leaveout ve Kfold çapraz doğrulama yöntemlerinde ve farklı komşuluk sayıları ve uzaklık bağıntılarında aynı sonuçlar elde edilmiştir.

4.6.2. Min-Maks Normalizasyon Yöntemi İle Seçilen Ortak Öznitelikler Ve Sınıflandırıcılardaki Etkinlikleri

Çizelge 4.13'te verilen sınıflandırma doğrulukları; LASSO, mRmR ve RELIEF yöntemleri ile seçilen 15'er öznitelik ile yapılan sınıflandırma sonuçlarının doğruluk değerlerinin altında kalmıştır. Dolayısıyla ortak öznitelikleri kullanılması verimli sonuçlar elde edilmesine katkı sağlamamıştır. Ancak üç öznitelik seçim yöntemi ile seçilen 15'er öznitelik ile yapılan sınıflandırma sonuçları, verilere Min-Maks normalizasyon yapılarak elde edilen sonuçlar olduğundan ortak öznitelikleri Min-Maks normalizasyon yaparak sınıflandırıcılarda kullanmak ve doğruluk değerlerini karşılaştırmak daha doğru bir yöntemdir. Bu sebeple Çizelge 4.14'te Min-Maks normalizasyon ile elde edilmiş olan

ortak öznitelikler verilmiştir. Bu özniteliklerle yapılan sınıflandırma sonuçları ayrıca değerlendirilmiştir.

Çizelge 4.14. Min-Maks Normalizasyon uygulanarak yapılan seçim işlemlerindeki ortak öznitelikler ve seçen yöntemlerin tablo şeklinde gösterimi.

	Öznitelik	Özniteliği seçen yöntem
1	Kalça açısının entropi değeri (42)	LASSO, mRmR ve RELIEF
2	Kalça açısının fast approximate entropi değeri (59)	LASSO, mRmR ve RELIEF
3	Diz açısının Higuchi değeri (76)	LASSO, mRmR ve RELIEF
4	Ayak bileği açısının multiscale entropi değeri (109)	LASSO, mRmR ve RELIEF
5	Ayak bileği açısının approximate entropi değeri (117)	LASSO, mRmR ve RELIEF
6	Kinetik verilerin orta değeri (9)	LASSO ve RELIEF
7	Kinetik verilerin harmonik ortalama değeri (20)	mRmR ve RELIEF
8	Kalça açısının Sample entropi değeri (48)	LASSO ve RELIEF
9	Diz açısının entropi değeri (72)	mRmR ve RELIEF
10	Ayak bileği açısının Katz değeri (107)	LASSO ve RELIEF

Çizelge 4.14'te görüldüğü üzere 3 öznitelik seçim yönteminde de çoğunlukla kinematik veriler ortak olarak seçilmiştir. Ortak olarak seçilen bu 10 öznitelik farklı kombinasyonlarla sınıflandırıcılarda çalıştırılmış ve sonuçları karşılaştırılmıştır. Bu kombinasyonlar:

- Grup 1: 42, 59, 76, 109 ve 117. Öznitelikler kullanılarak sınıflandırma,
- Grup 2: 9, 20, 48, 72 ve 107. Öznitelikler kullanılarak sınıflandırma,
- Grup 3: 9, 20, 42, 48, 59, 72, 76, 107, 109 ve 117. Öznitelikler kullanılarak sınıflandırma şeklindedir.

Grup 1,2 ve 3; LASSO, mRmR ve RELIEF yöntemlerinin üçünde de ortak seçilen öznitelikler, iki yöntemde seçilen öznitelikler ve en az iki seçim yönteminde seçilen öznitelikler olarak ifade edilebilir.

Çizelge 4.15. Min-Maks normalizasyon yöntemi ile LASSO, mRmR ve RELIEF yöntemlerinden seçilen ortak öznitelikler ve farklı sınıflandırıcılardaki performansları.

	Grup 1	Grup 2	Grup 3
SVM	100	96,88	100
KNN	100	100	100
Naive Bayes	93,75	84,38	87,50

Çizelge 4.15'te ortak seçilen özniteliklerden elde edilen sınıflandırma sonuçları bulunmaktadır. Tüm sınıf değişkenleri ve iki çapraz doğrulama yöntemi ile elde edilen farklı sonuçlar içinden en yüksek olanlar yazılmıştır. Çizelgeden de görüleceği üzere grup 1'deki 5 öznitelik (LASSO, mRmR ve RELIEF öznitelik seçme yöntemlerinin üçünde de seçilmiş olan) kullanılarak en yüksek sınıflandırıcı doğruluğuna ulaşılmıştır. Bu durum öznitelik seçimi ve bu seçim sonuçlarının birleştirilerek kullanılmasının sınıflandırma başarısına olumlu etki yaptığını göstermektedir.

BÖLÜM 5

SONUÇ ve TARTIŞMA

Parkinson hastalığı 1817 yılında James Parkinson tarafından keşfedildi ve o günden bu yana bilim insanlarının ilgisini çeken, güncelliğini halen koruyan bir hastalık olarak devam etti. Hastalığın halen bilinen kesin bir tedavi yönteminin olmaması; teşhisinin zor, gözleme dayalı ve maliyetli olması hem yeni teşhis yöntemleri geliştirmek hem de yeni tedavi yöntemleri geliştirmek için yeni araştırmalar yapılmasına sebep olmaktadır.

Hastalığın etkilediği önemli kısımlardan olan yürüyüş, geçmişten bu yana araştırmacıların yoğun ilgisini çekmiş bir konudur. Yürüyüş analizinin görecelilikten uzak, uygulanabilirliğinin kolay ve invaziv olmayan bir yöntem olması, hastalığın teşhisinin konforlu bir şekilde yapılabilmesine olanak sağlayabilir.

Bu tezde ‘Parkinson hastası ve sağlıklı bireylerden elde edilen yürüyüş verileri ile hastalığın teşhisi yapılabilir mi?’ sorusuna cevaben detaylı bir çalışma yapılmıştır. Bu tez çalışmasında makine öğrenmesi yöntemleri kullanılarak yalnızca yürüyüş verilerinden hasta-sağlıklı ayrımı yapabilen bir sistem geliştirilmeye çalışılmıştır.

Bu kapsamda makine öğrenmesi sınıflandırma yöntemlerinden SVM, KNN ve Naive Bayes olmak üzere üç farklı yöntem kullanılmış ve sonuçları karşılaştırılmıştır. Sınıflandırma performansını etkileyebileceği ön görülen farklı normalizasyon ve çapraz doğrulama yöntemleri de uygulanmış ve sınıflandırma performanslarına olan etkileri karşılaştırılarak yorumlanmıştır. Bu işlemlerin tamamı MATLAB programında yapılmış ve sonuçlar kaydedilmiştir.

Malezya MARA Teknoloji Üniversitesi'nden Prof. Nooritawati Md Tahir ve ekibinin topladığı yürüyüş verileri hem katılımcı sayısının az olması hem de detaylı yürüyüş öznitelikleri buldurumaması bakımından kısıtlıdır. Bu sebeple 21 olan öznitelik sayısı 152'ye yükseltilmiş ve bu 152 öznitelik ile normalizasyon, çapraz doğrulama, sınıflandırma ve öznitelik seçimi yapılmıştır. 21 öznitelik ile ve 152 öznitelik ile elde edilen sınıflandırma sonuçları karşılaştırıldığında 152 öznitelik ile yapılan sınıflandırmanın daha yüksek performans gösterdiği sonucuna varılmıştır.

Normalizasyon öncesi ve sonrasında yapılan sınıflandırma sonuçları karşılaştırılarak normalizasyon yöntemlerinin etkinliği irdelenmiştir. Min-Maks normalizasyonun sınıflandırma performansını sıfır ortalama ve normalizasyon yapılmayan durumlara göre önemli ölçüde arttırdığı gözlenmiştir. Ayrıca benzer sonuç öznitelik seçimi ile üç farklı yöntemden elde edilen, başarıya katkısının yüksek olduğu düşünülen ortak öznitelikler kullanılarak yapılan sınıflandırma sonuçlarında da gözlenmiştir. Seçilen ortak özniteliklere normalizasyon yapılmadığında elde edilen sınıflandırma sonuçlarının doğruluk oranı, Min-Maks normalizasyon ile elde edilen sınıflandırma sonuçlarının doğruluk oranının çok altında kalmıştır.

Sınıflandırma işlemlerinde eğitim ve test verilerinin farklı yaklaşımlarla bölünmesi prensibine dayanan iki farklı çapraz doğrulama yöntemi uygulanmıştır. Gerek 152 öznitelik ile yapılan sınıflandırmada gerekse öznitelik seçimi sonrası ve ortak özniteliklerle yapılan sınıflandırmalarda Leaveout ve Kfold çapraz doğrulama yöntemleri arasında anlamlı farklılıklar gözlenmemiştir. Ancak her ikisi de sınıflandırıcı performansını aynı derecede olumlu etkilemiştir.

Öznitelik seçme yöntemlerinden elde edilen sonuçlara göre hem normalizasyon yapılmayan durumda hem de Min-Maks normalizasyon yapılan durumda seçilen öznitelikler içinde kinematik verilerin sayısı dikkat çekmektedir. Normalizasyon yapılmadan LASSO, mRmR ve RELIEF yöntemleri ile öznitelik seçildiğinde toplamda 1 kinetik, 34 kinematik ve 10 temel yürüyüş özniteliği seçilmiştir. Min-Maks normalizasyon yapılarak öznitelikler seçildiğinde ise 3 yöntemde toplam 8 kinetik, 29 kinematik ve 4 temel yürüyüş özniteliği seçilmiştir. Bu sonuçlar bize yürüyüş ile Parkinson hastalığının tespitinde kinematik verilerin daha belirleyici olduğunu göstermektedir. Bu durum hastaların eklem ve kas katılıklarının olması sebebi ile hareket kabiliyetlerinin azalması yorumunu destekler.

3 öznitelik seçme yönteminden elde edilen 15'er öznitelik SVM, KNN ve Naive Bayes sınıflandırıcılarında ayrı ayrı çalıştırılmıştır. Öznitelik seçiminden gelen öznitelikler belirli bir artışla (biri birer, beşer beşer gibi) sınıflandırıcılarda çalıştırılır. Belirli bir doğruluk değerine kadar ilk özniteliklerden itibaren doğruluk grafiğinde bir artış beklenir ve bu artış doyum noktasına ulaşıncaya grafiğin sabitlenmesi beklenir. SVM sınıflandırıcıda RELIEF'ten seçilen öznitelikler kullanıldığında neredeyse düz bir eğri elde edilmiş, LASSO'dan seçilen öznitelikler kullanıldığında ise küçük değerlerde artış ve azalışlar olmuştur. Beklenen eğri SVM sınıflandırıcıda mRmR öznitelik seçim yönteminden gelen özniteliklerle yapılan sınıflandırmada çizilmiştir. Eğriliklerden bağımsız, tüm seçim yöntemlerinden gelen özniteliklerde yaklaşık 8 veya 9 öznitelikten itibaren %100 maksimum doğruluk değerine ulaşılmıştır.

KNN sınıflandırıcı modelinde RELIEF'ten ve LASSO'dan seçilen öznitelikler neredeyse düz bir eğri şeklinde doğruluk eğrisi çizerken, mRmR'da beklenen, artan bir doğruluk eğrisi gözlenmiştir. Eğriliklerden bağımsız olarak tüm seçim yöntemlerinden elde edilen sınıflandırma sonuçları yaklaşık 8 veya 9 öznitelikten itibaren %100 maksimum doğruluk değerine ulaşmıştır.

Naive Bayes sınıflandırıcı modelinde ise beklenen artışlı eğri mRmR seçiminden gelen özniteliklerle çizilmiştir. RELIEF'ten gelen öznitelikler, öznitelik sayısı arttıkça sınıflandırma doğruluğunu azaltırken, LASSO'dan gelen öznitelikler %87,50 ile %81,25 değerleri ile neredeyse düz bir doğruluk eğrisi çizdirmişlerdir. Naive Bayes sınıflandırma

yönteminde de elde edilebilecek en yüksek doğruluk değeri olan %100 değerine ulaşılmıştır.

Yukarıda bahsi geçen seçilen özneliklerle yapılan sınıflandırma sonuçları Min-Maks normalizasyon ile elde edilmiştir. Ayrıca Kfold ve Leaveout çapraz geçirme yöntemlerinin her ikisi de denendiğinde sonuçlarda anlamlı bir değişiklik gözlenmemiştir.

LASSO, mRmR ve RELIEF öznelik seçme yöntemlerinde seçilen öznelikler analiz edilmiş ve ortak olanlar kaydedilerek, ayrıca sınıflandırıcılarda çalıştırılmıştır. Özellikle Min-Maks normalizasyon uygulandığında elde edilen sonuçlarda doğruluk oranı %100'e ulaşmış ve 5 öznelik kullanıldığından sınıflandırıcının çalışma süresinde kısalmıştır.

Bu tez çalışmasından elde edilen sonuçlarla Parkinson hastalığının tespitinde yürüyüş verilerinin kullanımının önemli ve ayırt edici sonuçlar doğuracağı ispatlanmıştır. Hastalığın tespitinde yürüyüş verilerinin kullanılması hem hasta konforunu arttıracığı hem de hastane maliyetlerini düşüreceği öngörülmektedir. Hastalığın erken evrede kolay yöntemlerle tespiti, hastalığın ilerlemesinin durdurulması ve etkilerinin hafifletilmesini sağlayacağından ileri yaşlardaki hastaların yaşam kalitelerinin artmasını, hasta yakınlarının da daha rahat bir hayat sürmelerini sağlayacaktır. Hastalığın genç yaşlarda görüldüğü "genç başlangıçlı Parkinson" olarak isimlendirilen hastaların da uygun tedavi yöntemlerinin belirlenmesi açısından hastalığın erken tespiti değerlidir. Yalnızca yürüyüş verileri ile erken tespit yapabilen bir sistemin kurulması hekimlerin elini güçlendirecek ve teşhis sürecini hızlandıracaktır. Bu durum hem hekim hem hastane hem hasta hem de hasta yakınları açısından avantajdır.

İlerleyen süreçlerde daha fazla hasta ve sağlıklı bireyin katılımı ile daha büyük bir veri seti ile farklı sınıflandırma yöntemleri kullanılarak daha yüksek performansa sahip tasarımlar hakkında yeni çalışmalar yapılabilir. Yürüyüşten elde edilen öznelik sayısı artırılarak daha ayrıntılı sonuçların elde edilmesi sağlanabilir. Yürüyüşe ek olarak denge ile ilgili kayıtların kullanılması da çalışmaların kapsamını genişletebilir. Aynı zamanda yalnızca hastalık tespiti yapan çalışmalar yerine evre tayini de yapabilen hibrit sistemler geliştirilebilir.

Sistem mobilleştirilerek bir uygulama haline getirilebilir, uygulamayı kullananlar sadece yürüyerek Parkinson hastası olma ya da olmama ihtimallerini öğrenebilirler. Parkinson hastası olma ihtimali olan kullanıcının bir nöroloğa muayene olması yönünde uyarılar yapılarak erken evrede hastaların hekime başvurmaları sağlanabilir.

Daha büyük veri setlerinin oluşturulacağı arařtırmalar ile derin öğrenme yöntemleri de çalışmalara dahil edilebilir. Derin öğrenme günümüzün güncel çalışma alanlarındandır. Özellikle büyük ve karmaşık veri setleri için daha hassas ve kesin sonuçlar elde edilmesini sağlamaktadır.



KAYNAKLAR

Abdulhay, E., Arunkumar, N., Narasimhan, K., Vellaiappan, E., & Venkatraman, V. (2018). Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease. *Future Generation Computer Systems*, 83, 366-373.

Acharya, R. U., Lim, C. M., & Joseph, P. (2002). Heart rate variability analysis using correlation dimension and detrended fluctuation analysis. *Itbm-Rbm*, 23(6), 333-339.

Adams, W. R. (2018). The detection of hand tremor through the characteristics of finger movement while typing. *bioRxiv*, 385286.

Akşehirli, Ö. Y., Ankaralı, H., Aydın, D., & Saraçlı, Ö. (2013). Tıbbi Tahminde Alternatif Bir Yaklaşım: Destek Vektör Makineleri. *Türkiye Klinikleri Journal of Biostatistics*, 5(1).

Alsancak, S. (2015). Yürüyüş terminolojisi. *Ankara Sağlık Hizmetleri Dergisi*, 14(2), 1-6.

Altun, A. M., Ozbek, S. E., Zarifoglu, M., & Özkaya, G. (2013). Parkinson Hastalığında yürüme ve dengenin değerlendirilmesi. *Parkinson Hastalığı ve Hareket Bozukluğu Dergisi*, 16, 1-8.

Altunkaynak, A., Başakın, E. E., & Kartal, E. (2020). Dalgacık K-En Yakın Komşuluk Yöntemi İle Hava Kirliliği Tahmini. *Uludağ University Journal of The Faculty of Engineering*, 25(3), 1547-1556.

Alty, J. E., Cosgrove, J., Lones, M. A., Smith, S. L., Possin, K., Schuff, N., & Jamieson, S. (2016). Clinically 'slight'bradykinesia in Parkinson's disease is accurately detected using evolutionary computation analysis of finger tapping. *Mov Disord*, 31.

Arora, S., Baig, F., Lo, C., Barber, T. R., Lawton, M. A., Zhan, A., Rolinski, M., Ruffmann, C., Klein, J., Rumbold, J., Louvel, A., Zaiwalla, Z., Lennox, G., Quinnell, T., Dennis, G., Wade-Martins, R., Ben-Shlomo, Y., Little, M. & Hu, M. T. (2018). Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. *Neurology*, 91(16), e1528-e1538.

Ayhan, S., & Erdoğan, Ş. (2014). Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 9(1), 175-201.

Balaji, E., Brindha, D., & Balakrishnan, R. (2020). Supervised machine learning based gait classification system for early detection and stage classification of Parkinson's disease. *Applied Soft Computing*, 94, 106494.

Baydın, Ş. S. (2021) Parkinson Hastalığında Derin Beyin Stimülasyonu: Nöroradyoloji ve Anatomik Hedefler. *Türk Nöroşir Derg* 31(3):267-270, 2021

Błaszczak, J. W., Orawiec, R., Duda-Kłodowska, D., & Opala, G. (2007). Assessment of postural instability in patients with Parkinson's disease. *Experimental Brain Research*, 183(1), 107-114.

Cavagna, G. A., & Margaria, R. (1966). Mechanics of walking. *Journal of applied physiology*, 21(1), 271-278.

Chou, C., Chou, S. H., Chen, Y. C., & Yang, C. J. (2020). Using machine learning methods to detect physical conditions with postural balance. *Journal of Ambient Intelligence and Humanized Computing*, 1-7.

Combs, S. A., Diehl, M. D., Filip, J., & Long, E. (2014). Short-distance walking speed tests in people with Parkinson disease: reliability, responsiveness, and validity. *Gait & posture*, 39(2), 784-788.

Costa, M., Goldberger, A. L., & Peng, C. K. (2005). Multiscale entropy analysis of biological signals. *Physical review E*, 71(2), 021906.

Dalyan, T. (2006). Makine öğrenmesinde 1R algoritması ve ikinci kuralın (2R) oluşturulması (Master's thesis, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü).

de Pina Filho, A. C., Dutra, M. S., & Raptopoulos, L. S. (2005). Modeling of a bipedal robot using mutually coupled Rayleigh oscillators. *Biological cybernetics*, 92(1), 1.

Di Monte, D. A., Lavasani, M., & Manning-Bog, A. B. (2002). Environmental factors in Parkinson's disease. *Neurotoxicology*, 23(4-5), 487-502.

Faul, S. (2021). Kolmogorov Complexity (<https://www.mathworks.com/matlabcentral/fileexchange/6886-kolmogorov-complexity>), MATLAB Central File Exchange. Retrieved December 20, 2021.

Fonti, V., & Belitser, E. (2017). Feature selection using lasso. *VU Amsterdam Research Paper in Business Analytics*, 30, 1-25.

Göz, E., ÇOLAKOĞLU, B. D., ÇAKMUR, R., & BALCI, B. (2020). Erken dönem parkinson hastalarında pilates ve elastik bantlama uygulamalarının denge ve postüral kontrol üzerine etkisi (*Doctoral dissertation, Doktora Tezi, Dokuz Eylül Üniversitesi Sağlık Bilimleri Enstitüsü, İzmir*).

Gözellik, E. Ö. (2018). 16.12.2021 tarihinde <https://www.aa.com.tr/tr/saglik/dunyada-7-milyondan-fazla-parkinson-hastasi-bulunuyor/1113515> adresinden erişildi.

Gülgezen, G., Çataltepe, Z., & Yu, L. MRMR Algoritması Kullanılarak Kararlı Öznitelik Seçimi Stable Feature Selection Using MRMR Algorithm.

Hasomed, (2021). 18.12.2021 tarihinde <https://hasomed.de/en/products/rehagait/> adresinden erişildi.

Hayot, C., Sakka, S., & Lacouture, P. (2011). Kinematic and dynamic analysis of gait determinants theory. *Computer Methods in Biomechanics and Biomedical Engineering*, 14(1), 153-154.

Hemensaglik, (2021). 16.12.2021 tarihinde <https://www.hemensaglik.com/makale/parkinson-teshisinin-konulmasi> adresinden erişildi.

Hubble, R. P., Naughton, G. A., Silburn, P. A., & Cole, M. H. (2015). Wearable sensor use for assessing standing balance and walking stability in people with Parkinson's disease: a systematic review. *PloS one*, 10(4), e0123705.

Impedovo, D., Pirlo, G., & Vessio, G. (2018). Dynamic handwriting analysis for supporting earlier Parkinson's disease diagnosis. *Information*, 9(10), 247.

Kasap, M., & Akpınar, G. (2011). Parkinson Hastalığı Tanısında Proteomiks Yaklaşımlarının Yeri ve Biyobelirteç Arayışları/Proteomics Approaches to Parkinson's Disease and Search for Potential Biomarkers. *Türkiye Klinikleri. Tıp Bilimleri Dergisi*, 31(3), 691.

Kaynar, O., Arslan, H., Görmez, Y., & Demirkoparan, F. (2017, September). Feature selection methods in sentiment analysis. In *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1-5). IEEE.

Kimmeskamp, S., & Hennig, E. M. (2001). Heel to toe motion characteristics in Parkinson patients during free walking. *Clinical biomechanics*, 16(9), 806-812.

Kurt, İ. (2019). Makine öğrenmesi yöntemleri kullanılarak sesin müzikal öznitelikleri ile Parkinson hastalığının tespiti (*Master's thesis, Trakya Üniversitesi Fen Bilimleri Enstitüsü*).

Kuo, A. D. (2007). The six determinants of gait and the inverted pendulum analogy: A dynamic walking perspective. *Human movement science*, 26(4), 617-656.

Lee, K.(2021). Fast Approximate Entropy (<https://www.mathworks.com/matlabcentral/fileexchange/32427-fast-approximate-entropy>), *MATLAB Central File Exchange*. Retrieved December 21, 2021.

Little, M., McSharry, P., Hunter, E., Spielman, J., & Ramig, L. (2008). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *Nature Precedings*, 1-1.

Magris, M. (2021). Detrended fluctuation analysis (DFA) (<https://www.mathworks.com/matlabcentral/fileexchange/67889-detrended-fluctuation-analysis-dfa>), *MATLAB Central File Exchange*. Retrieved December 21, 2021.

Malik, J. (2021). Multiscale Sample Entropy (<https://www.mathworks.com/matlabcentral/fileexchange/62706-multiscale-sample-entropy>), *MATLAB Central File Exchange*. Retrieved December 20, 2021.

Manap, H. H., Tahir, N. M., Yassin, A. I. M., & Abdullah, R. (2011a, June). Anomaly gait classification of parkinson disease based on ann. In *2011 IEEE International Conference on System Engineering and Technology* (pp. 5-9). IEEE.

Manap, H. H., Tahir, N. M., & Yassin, A. I. M. (2011b, December). Statistical analysis of parkinson disease gait classification using Artificial Neural Network. In *2011 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)* (pp. 060-065). IEEE.

Mar-systems, (2021). 18.12.2021 tarihinde https://www.mar-systems.co.uk/motionmetrix_3d-running-analysis/reports/ adresinden erişildi.

Mc Ardle, R., Del Din, S., Galna, B., Thomas, A., & Rochester, L. (2020). Differentiating dementia disease subtypes with gait analysis: feasibility of wearable sensors?. *Gait & posture*, 76, 372-376.

Mete, S., ÇAKIR, O., Bayat, O., Duru, D. G., & Duru, A. D. (2020). Gözbebeği hareketleri temelli duygu durumu sınıflandırılması. *Bilişim Teknolojileri Dergisi*, 13(2), 137-144.

Medicaptureurs, (2021). 18.12.2021 tarihinde <https://www.medicaptureurs.com/produits/winshoe-2/> adresinden erişildi.

Monge-Álvarez, J. (2021). Higuchi and Katz fractal dimension measures (<https://www.mathworks.com/matlabcentral/fileexchange/50290-higuchi-and-katz-fractal-dimension-measures>), *MATLAB Central File Exchange*. Retrieved December 20, 2021.

Novel, (2021). 18.12.2021 tarihinde https://www.novel.de/products/pedar/pedar-software/#pedar_s adresinden erişildi.

Oliveira, L. M., Gasser, T., Edwards, R., Zweckstetter, M., Melki, R., Stefanis, L., ... & Outeiro, T. F. (2021). Alpha-synuclein research: defining strategic moves in the battle against Parkinson's disease. *npj Parkinson's Disease*, 7(1), 1-23.

Orakçı, M., Ciylan, B., Kök, İ., & Sevri, M. (2019). Suç Analizinde Veri Madenciliği Teknikleri Ve Makine Öğrenmesi Algoritmalarının Kullanılması.

Özer, A. F., Kaner, T., & Bozdoğan, C. (2014). Sagittal Balance in the Spine. *Turkish Neurosurgery*, 24(Sup1), 13-19.

Parnandi, A. (2021). Approximate Entropy (<https://www.mathworks.com/matlabcentral/fileexchange/26546-approximate-entropy>), *MATLAB Central File Exchange*. Retrieved December 21, 2021.

Pehlivan, N., Tunca, C., Salur, G., & Ersoy, C. (2017, May). Gait analysis using kinect: Towards in-home gait analysis. *In 2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6), H2039-H2049.

Shetty, S., & Rao, Y. S. (2016, August). SVM based machine learning approach to identify Parkinson's disease using gait analysis. *In 2016 International Conference on Inventive Computation Technologies (ICICT)* (Vol. 2, pp. 1-5). IEEE.

Silahtaroğlu, G. (2016). *Veri Madenciliği Kavram ve Algoritmaları*. İstanbul:Papatya.

Soyuer, F., & İsmailoğulları, S. (2009). Yaşlılık ve denge. *Türk Beyin Damar Hastalıkları Dergisi*, 15(1), 1-5.

Tahir, N. M., & Manap, H. H. (2012). Parkinson disease gait classification based on machine learning approach. *Journal of Applied Sciences(Faisalabad)*, 12(2), 180-185.

Taşcı, E., & Onan, A. (2016). K-en yakın komşu algoritması parametrelerinin sınıflandırma performansı üzerine etkisinin incelenmesi. *Akademik Bilişim*, 1-8.

UNO Biomechanics (2021). UNO Biomechanics Nonlinear Analysis Toolbox (<https://www.mathworks.com/matlabcentral/fileexchange/71907-uno-biomechanics-nonlinear-analysis-toolbox>), *MATLAB Central File Exchange*. Retrieved December 20, 2021.

Veribilimiokulu, (2021). 22.12.2021 tarihinde <https://www.veribilimiokulu.com/bir-bakista-k-fold-cross-validation/> adresinden erişildi.

Yaliman, A., & Ően, E. İ. (2011). Parkinson Hastalıđı ve Rehabilitasyonu. *Turkish Journal of Physical Medicine & Rehabilitation/Turkiye Fiziksel Tip ve Rehabilitasyon Dergisi*, 57(1).

Yavuzer, G. (2014). Yürüme analizi ve temel kavramlar. *Türk Ortopedi ve Travmatoloji Birliđi Derneđi Dergisi*, İstanbul, 13, 304-308.

YIGIT, G., & ARICIOGLU, F. (2015) Günümüz ve Gelecekte Parkinson Hastalıđı için Farmakolojik Tedavi Yaklaşımaları. *Clinical and Experimental Health Sciences*, 5(4), 265-273.



ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı – Soyadı: Sinem AYNA ALTUNTAŞ

EĞİTİM BİLGİLERİ

2018 - 2022	Trakya Üniversitesi (Yüksek Lisans) Hesaplamalı Bilimler
2010 – 2014	Erciyes Üniversitesi Biyomedikal Mühendisliği
2005 – 2009	Alparslan Anadolu Lisesi Fen Bilimleri

İŞ TECRÜBELERİ

- 2015-2016--MED-SAN TIBBİ CİHAZLAR/ANKARA- Sorumlu Müdür
- 2017-2018--ATESE SAĞLIK HİZMETLERİ/ANKARA- Biyomedikal Mühendisi
- 2018-Halen--TRAKYA ÜNİVERSİTESİ-EDİRNE-Öğretim Görevlisi