T.C.

TRAKYA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK PARMAK VURUŞ HAREKETİ VERİLERİ İLE PARKİNSON HASTALIĞI TESPİTİ

BAHAR DEMİR

YÜKSEK LİSANS TEZİ

HESAPLAMALI BİLİMLER ANABİLİM DALI

Tez Danışmanı: Doç. Dr. OĞUZHAN ERDEM

EDİRNE-2022

BAHAR DEMİR'in hazırladığı "**MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK PARMAK VURUŞ HAREKETİ VERİLERİ İLE PARKİNSON HASTALIĞI TESPİTİ**" başlıklı bu tez, tarafımızca okunmuş, kapsam ve niteliği açısından **HESAPLAMALI BİLİMLER** Anabilim Dalında bir **Yüksek lisans tezi** olarak kabul edilmiştir.

Jüri Üyeleri:	İmza
Doç. Dr. Oğuzhan ERDEM	……
Dr. Öğr. Üyesi Aydın CARUS	……
Dr. Öğr. Üyesi Barış ÖZYER	……

Tez Savunma Tarihi: 17/01/2022

Bu tezin Yüksek Lisans tezi olarak gerekli şartları sağladığını 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 çalışmasında, 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

Bahar DEMİR

İmza

Yüksek Lisans Tezi

Makine Öğrenimi Teknikleri Kullanılarak Parmak Vuruş Hareketi Verileri ile Parkinson Hastalığı Tespiti

T.Ü. Fen Bilimleri Enstitüsü

Hesaplamalı Bilimler Anabilim Dalı

ÖZET

Parkinson hastalığı dünyada en yaygın görülen ikinci nörodejeneratif (sinir hasarına bağlı geri dönüşü olmayan) rahatsızlıktır ve henüz kesin bir tedavisi bulunmamaktadır. Hastanın motor becerilerinde meydana gelen ve günlük aktivitelerini olumsuz yönde etkileyen hareket bozuklukları (yürüme, konuşma, denge vb.), uzman klinisyenler tarafından ilaç tedavileri uygulanarak azaltılmaya çalışılmaktadır.

El hareketlerindeki bozulmalar (titreme, hareketlerde yavaşlama vb.) hastalığın ilk belirtileri arasında yer almakta olup, hastalığın erken evrede tespiti ile uygulanacak ilaç tedavileri maddi ve manevi açıdan büyük önem taşımaktadır. Bu amaçla, bu çalışmada Parkinson hastalığında klavye tuş vuruş dinamikleri 14 farklı veri kümesinde ilaç kullanım durumlarına, hastalık şiddetine ve cinsiyete göre incelenmiştir. Parkinson hastası ve sağlıklı katılımcılardan oluşan, donma, gecikme ve uçuş sürelerini içeren gün içerisinde arka planda çalışıp kayıt altına alınan ham verilerden öznitelikler çıkartılmış, makine öğrenmesi algoritmaları uygulanarak hangi klavye tuş vuruş özelliklerinin Parkinson hastalarını sağlıklı bireylerden ayırmada daha başarılı olduğu araştırılmıştır.

Yıl : 2022

Sayfa Sayısı : 84

Anahtar Kelimeler : Parkinson Hastalığı, Parmak Vuruşu, Makine Öğrenmesi, Klavye Tuş Vuruşu, El-Parmak Hareketleri Analizi, Öznitelik Seçimi, Sınıflandırma

Master’s Degree Thesis
Detection of Parkinson's Disease with Finger Tapping Data Using Machine Learning Techniques
Trakya University Institute of Natural and Applied Sciences
Computational Science Department

ABSTRACT

Parkinson's disease is the second most common neurodegenerative (irreversible due to nerve damage) disorder in the world and there is no definitive treatment yet. Movement disorders (gait, speech, balance, etc.), are tried to be reduced by the application of drug treatments by specialist clinicians which occur in the motor skills of the patient and negatively affect his daily activities.

Impairments in hand movements (tremor, bradykinesia, etc.) are among the first symptoms of the disease and drug treatments to be applied with the detection of the disease at an early stage are of great importance in material and spiritual terms. For this purpose, in this study keyboard keystroke dynamics in Parkinson's disease were investigated in 14 different datasets according to drug use status, disease severity and gender. Features were extracted from the raw data including hold, latency and flight times recorded in the background during the day from Parkinson's patients and healthy participants and by applying machine learning algorithms, which keyboard keystroke characteristics were more successful in distinguishing Parkinson's patients from healthy individuals were investigated.

Year : 2022

Number of Pages : 84

Keywords : Parkinson's Disease, Finger Tapping, Machine Learning, Keystroke, Hand-Finger Movement Analysis, Feature Selection, Classification

ÖNSÖZ

Bu çalışma, Trakya Üniversitesi Fen Bilimleri Enstitüsü Hesaplamalı Bilimler Anabilim Dalı'nda yapılan "Makine Öğrenimi Teknikleri Kullanılarak Parmak Vuruş Hareketi Verileri ile Parkinson Hastalığı Tespiti" isimli yüksek lisans tez çalışmasını içermektedir.

Yüksek lisans öğrenimime en büyük adımı atmamı ve her zaman daha da ilerisini görmemi sağlayan, yardım ve desteklerini benden hiçbir zaman esirgemeyen değerli danışman hocam, Sayın Doç. Dr. Oğuzhan ERDEM'e,

Çalışmamın tüm aşamalarında engin tecrübe ve bilgisini benimle paylaşarak, emin adımlarla ilerlememi sağlayan Sayın Dr. Öğr. Üyesi Sezer ULUKAYA'ya,

Her koşulda ve her zaman yanımda olan, desteklerini hiçbir zaman esirgemeyen, zor zamanlarda dahi çözümler üreten değerli arkadaşlarım İlke KURT ile Sinem AYNA ALTUNTAŞ'a ve canım aileme teşekkür ve sevgilerimi sunarım.

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vi**

İÇİNDEKİLER ..... **vii**

ŞEKİLLER DİZİNİ ..... **x**

ÇİZELGELER DİZİNİ ..... **xii**

SİMGELER VE KISALTMALAR DİZİNİ ..... **xiii**

BÖLÜM 1 ..... **1**

GİRİŞ **1**

BÖLÜM 2 ..... **4**

PARKİNSON HASTALIĞI VE EL HAREKETLERİ **4**

2.1. Parkinson Hastalığı 4
2.1.1. Hastalığın Nedenleri 5
2.1.2. Hastalığın Belirtileri 5
2.1.3. Parkinson Hastalığı Tespitinde Kullanılan Yöntemler 7
2.2. El Hareketleri 8
2.3. El Hareketlerinde Bozulmalar ve Parkinson Hastalığı 9
2.3.1. Klavye Tuş Vuruş Dinamikleri 9
2.3.2. Parmak Vuruşu Hareketi (Finger Tapping) 9
2.3.3. El Açma-Kapama Hareketi (Opening-Closing) 9

2.3.4. El Döndürme Hareketi (Pronation-Supination) 10
2.4. El Hareketleri ile Yapılan Parkinson Hastalığı Çalışmaları 10

BÖLÜM 3 15

MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK PARMAK VURUŞ HAREKETİ VERİLERİ İLE PARKİNSON HASTALIĞI TESPİTİ 15

3.1. Makine Öğrenimi 15
3.2 Verilerin Elde Edilmesi 16
3.3. Öznitelik Çıkarımı 18
3.3.1. Minimum 19
3.3.2. Maksimum 20
3.3.3. Ortalama 20
3.3.4. Medyan 20
3.3.5. Standart Sapma 20
3.3.6. Aralık 20
3.3.7. Entropi 21
3.3.8. Basıklık 21
3.3.9. Çarpıklık 21
3.3.10. Kolmogorov Karmaşıklığı 21
3.3.11. Lempel-Ziv Karmaşıklığı 21
3.3.12. Örnek Entropi 22
3.3.13. Higuchi Fraktal Boyut Analizi 22
3.3.14. Katz Fraktal Boyut Analizi 22
3.4. Normalizasyon Yöntemleri 22
3.4.1. Minimum – Maksimum Normalizasyon 23
3.4.2. Z-skor Normalizasyon 23
3.5. Öznitelik Seçim Yöntemleri 23
3.5.1. Minimum Artıklık Maksimum Bağlantı [Minimum Redundancy Maximum Relevance (mRmR)] 24
3.5.2. RELIEF 25
3.5.3. Ardışık İleri Yönde Seçim [Sequential Forward Feature Selection (SFS)] 25
3.5.4. Rastgele Orman 25
3.6. Sınıflandırma Yöntemleri 26
3.6.1. k-En Yakın Komşu Algoritması 26
3.6.2. Destek Vektör Makineleri 27

3.6.2. Rastgele Orman 29
3.7. Birini Dışarda Bırakma [Leave-one-out (LOO)] Çapraz Doğrulama Yöntemi ... 29
3.8. Değerlendirme Ölçütleri 30
3.8.1. Doğruluk 31
3.8.2. Duyarlılık 31
3.8.3. Kesinlik 31
3.8.4. F-ölçütü 31
3.8.5. Matthews Korelasyon Katsayısı 32

BÖLÜM 4 33

DENEYSEL SONUÇLAR 33

4.1. Veri Kümesi 34
4.2. Normalizasyon Yöntemlerinin Etkinliği 35
4.3. Öznitelik Seçim Yöntemlerinin Etkinliği 40
4.4. Ortak Seçilen Özniteliklerle Performans Değerlendirmesi 46

BÖLÜM 5 54

SONUÇ VE TARTIŞMA 54

KAYNAKLAR 58

EKLER 63

ÖZGEÇMİŞ 70

ŞEKİLLER DİZİNİ

Şekil 2.1. Substantia Nigra 5
Şekil 2.2. Basit El Anatomisi 8
Şekil 2.3. Klavye tuş vuruşlarının tarafları (Sarı alan sol tuşları; mavi alan sağ tuşları temsil etmektedir) 10
Şekil 3.1. Genel akış şeması 16
Şekil 3.2. Katılımcı Sağlık Bilgileri 17
Şekil 3.3. Katılımcı Tuş Vuruş Verileri 18
Şekil 3.4. Sarıcı öznitelik seçme yöntemi temsili 24
Şekil 3.5. k-En Yakın Komşu Algoritması ile k=3 komşuluğunda verinin sınıflandırılması 27
Şekil 3.6. DVM hiper düzlem ve destek vektörlerin belirlenmesi 28
Şekil 3.7. DVM çekirdek fonksiyonları 29
Şekil 3.8. Birini dışarda bırakma çapraz doğrulama yöntemi 30
Şekil 4.1. Tüm Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları 38
Şekil 4.2. Hafif derece Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları 39
Şekil 4.3. Hastalık şiddetine göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları 39
Şekil 4.4. Cinsiyete göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları 40
Şekil 4.5. Tüm Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları 43

Şekil 4.6. Hafif derece Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları 44
Şekil 4.7. Hastalık şiddetine göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları 45
Şekil 4.8. Cinsiyete göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları 45
Şekil 4.9. Tüm Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları 49
Şekil 4.10. Hafif derece Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları 50
Şekil 4.11. Hastalık şiddetine göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları 51
Şekil 4.12. Cinsiyete göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları 51

ÇİZELGELER DİZİNİ

Çizelge 2.1. Klavye tuş vuruş yönleri. 11
Çizelge 2.2. Literatür çalışmaları. 14
Çizelge 3.1. Veri seti özellikleri 17
Çizelge 3.2. Çıkarılan Öznitelikler 19
Çizelge 3.3. Karmaşıklık matrisi 31
Çizelge 4.1. Oluşturulan veri kümeleri 35
Çizelge 4.2. Normalizasyon yöntemlerine göre sınıflandırma başarı performansları 36
Çizelge 4.3. Öznitelik seçim yöntemleri ile seçilen ilk 20 öznitelik 41
Çizelge 4.4. Öznitelik seçim yöntemlerine göre sınıflandırma başarı performansları 42
Çizelge 4.5. Ortak seçilen ilk 5 öznitelik 47
Çizelge 4.6. Orta seçilen özniteliklere göre sınıflandırma başarı performansları 48
Çizelge 4.7. Veri kümelerine ait sınıflandırma doğruluk değerleri 53

SİMGELER VE KISALTMALAR DİZİNİ

Simgeler	**Açıklamalar**
$\bar{x}$	: Ortalama
s	: Standart Sapma

Kısaltmalar	**Açıklamalar**
AUC	: ROC Eğrisi Altında Kalan Alan (Area Under the Curve)
BDB	: Birini Dışarda Bırakma
BPHDÖ	: Birleşik Parkinson Hastalığı Değerlendirme Ölçeği
DVM	: Destek Vektör Makineleri
GN	: Gerçek Negatif
GP	: Gerçek Pozitif
HFBA	: Higuchi Fraktal Boyut Analizi
k-EYK	: k-En Yakın Komşu
KFBA	: Katz Fraktal Boyut Analizi
mRmR	: Minimum Artıklık Maksimum Bağlantı
PH	: Parkinson Hastalığı
RO	: Rastgele Orman

SFS	: Ardışık İleri Yönde Seçim (Sequential Forward Feature Seleciton)
YN	: Yanlış Negatif
YP	: Yanlış Pozitif

BÖLÜM 1

GİRİŞ

Parkinson hastalığı (PH), dünya çapında yaygın olarak, sinir hasarları ve ölümleri sonucu meydana gelen bir hastalıktır (Rosas vd., 2016). İlk olarak 1817'de James Parkinson tarafından "titreşimli felç" (vibratile paralysis) adıyla tanımlanmıştır (Yılancıoğlu, 2017). Ancak PH'ı nörolojik bir sendrom olarak gören James Parkinson'dan yıllar önce, 18. yy'da dinlenme titremesi ile ilgili çalışmalar bulunmaktadır (Stern, 1989). Ayrıca hastalığın varlığı hakkındaki bilgilere M.Ö. 1000'den kalma geleneksel Hint metinlerinde ve eski Çin kaynaklarında da rastlanmaktadır (Goetz, 2011; Manyam, 1990).

Merkezi sinir sisteminin substantia nigra bölgesinde, dopaminin azalması sonucu meydana gelen PH'ta, dopamin miktarındaki değişim ile kas kontrol mekanizması etkilenmektedir (Xia, Yao, Ye, & Cheng, 2019). Hastalığın ilk evresinde el ve kollarda titreme ile başlarken, ilerleyen evrelerde kaslarda katılık ve hareketlerde yavaşlama meydana gelmektedir. Hastalığın kesin bir tedavisi yoktur (Tamson, 2018). Uygulanan tedavi ve yöntemler, hastalığın ilerlemesiyle oluşan belirtileri ortadan kaldırarak günlük yaşam şartlarını iyileştirmeye yöneliktir (Zhang vd., 2020).

Uzman klinisyenler, hastaların aile hikayelerine bakarak ve nörogörüntüleme teknikleri gibi tanı yöntemlerini kullanarak hastalığın teşhisini yapar ve şiddetini belirlerler. Ancak hastalığın ilk belirtileri hafif olduğundan diğer nörolojik rahatsızlıklarla

karışıtırılabilmektedir. Yanlış tanı ile kullanılan ilaçlar hastalığın belirtilerini azaltmak yerine şiddetini arttırma riski oluşturmaktadır. Bu nedenlerden dolayı, hatalığın erken evrede teşhis edilmesi ile uygulanacak tedavi yöntemlerinin doğru bir şekilde belirlenmesi hastalığın ilerleyişi açısından oldukça önemlidir.

El hareketleri, günlük yaşamda bireyin temel gereksinimlerini karşılayabilmesi için gereklidir. El ve kollarda meydana gelen bozulmalar hareketin yapılmasını engellerken bireyin sosyal yaşantısını da etkileyebilmektedir. Parkinson hastalarında, dopamin seviyesinde meydana gelen azalmadan dolayı sinirler arası mesaj iletiminde sorunlar meydana gelir. İlk evrelerde tek taraflı olarak seyreden titreme, hareketlerde yavaşlama, kaslarda katılık ve donma gibi belirtiler tedavi edilmediği takdirde diğer tarafı da etkileyebilmektedir (Perlmutter, 2009).

El hareketleri ile yapılan çalışmaların bazılarında işaret ve baş parmağı birbirine vurma, el açıp kapama, el döndürme testleri bulunmaktadır. Bu testler ile hız, açı ve süre gibi özellikler hesaplanarak PH ile sağlıklı birey ayrımı yapılmıştır (Butt vd., 2018; Kandori vd. 2004; Gao vd. 2018).

Bireyin günlük klavye kullanım davranışı ile de hesaplanabilen donma süresi, gecikme süresi ve uçuş süresi gibi özellikler Parkinson hastalarını sağlıklı bireylerden ayırmada yüksek başarı oranına sahip yöntemlerden biridir. Donma süresi, bireyin o anki tuşa basma ve bırakma arasında geçen süre; gecikme süresi, ard arda gelen tuşlara basma anları arasında geçen süre; uçuş süresi ise ard arda gelen tuşlardan ilkini bırakıp ikinci tuşa basma arasında geçen süre olmakla beraber bu tuş vuruş dinamikleri kaslarda meydana gelen titreme, hareketlerde yavaşlama, katılık ve donma gibi özellikleri süre bazında tespit edebilmektedir (Schaap, 2018).

Klavye tuş vuruş dinamikleri, yazma davranışı gibi kişiden kişiye değişiklik gösteren özelliklere sahiptir. Tuş vuruşu ile yapılan daha önceki çalışmalarda harflere basma süreleri incelenmiştir. Ancak hangi harfe basıldığı bilgisi mevcut olmayan çalışmalarda tuşun sol ya da sağ tarafta yer alan bir tuş olduğu bilinmektedir. Boşluk tuşuna iki kere basma verileri ise bilindiği halde bazı çalışmalarda dahil edilmemiştir (Adams, 2017; Tamson, 2018; Schaap, 2018).

Bu tez çalışmasında, Physionet veri tabanından Parkinson hastası ve sağlıklı bireylere ait tuş vuruşu süresi ham verileri kullanılmıştır (Goldberger vd., 2000). Veri seti

her bireyin günlük tuş vuruşlarını içerdiğinden dolayı veri uzunlukları dengesizdi. Bundan dolayı her bireyden tek günlük, en az 2000 tuş vuruşu bulunan veriler çalışmaya dahil edilmiştir. Donma, gecikme ve uçuş sürelerini içeren ham veriler gruplandırılarak 378 öznitelik çıkartılmıştır. Elde edilen bu öznitelikler ile k-En Yakın Komşu Algoritması (k-EYK), Destek Vektör Makinaları (DVM) ve Rastgele Orman Algoritması ile hasta-sağlıklı ayrımı performansları incelenmiştir. Özniteliklerin en belirleyici olanlarını ayırt edebilmek için Minimum Artıklık Maksimum Bağlantı (mRmR – Minimum Redundancy Maximum Relevance), Relief, Ardışık İleri Yönde Seçim (SFS- Sequential Forward Feature Selection) ve Rastgele Orman (Random Forest) öznitelik seçilim yöntemleri uygulanmıştır. Elde edilen öznitelikler ile sınıflandırıcı performansları tekrar hesaplanarak önceki sonuçlarla kıyaslanmıştır. Veri kümesi, katılımcıların cinsiyet, hastalık şiddeti ve ilaç kullanım durumları gibi 14 farklı veri kümesine ayrılmış ve tüm işlemler her veri kümesi için tekrarlanmıştır.

Bu tez çalışmasında Bölüm 2, çalışmanın ana konusu PH'ı, el hareketlerini ve PH ile el hareketleri arasındaki bağlantıyı açıklamaktadır. Makine öğrenmesi algoritmaları, öznitelik çıkartma ve öznitelik seçme algoritmaları hakkında bilgiler Bölüm 3'te yer almaktadır. Bölüm 4'te yapılan deneyler ile elde edilen sonuçlar incelenmiştir. Sonuçlar ve yapılacak çalışmalara katkısı Bölüm 5'te belirtilmiştir.

BÖLÜM 2

PARKİNSON HASTALIĞI VE EL HAREKETLERİ

2.1. Parkinson Hastalığı

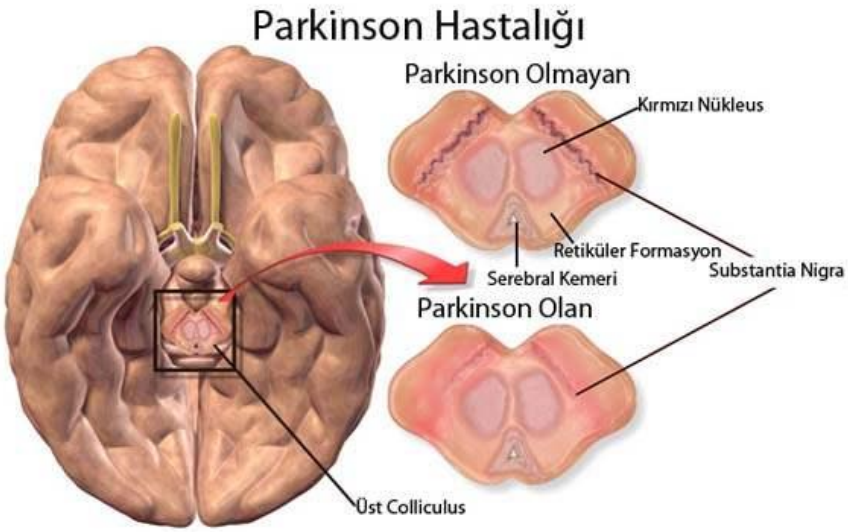
PH, dünya nüfusunda 6,3 milyon insanı etkileyen ilerleyici nörodejeneratif bir rahatsızlıktır (Adams, 2018). Nedeni bilinmemekle birlikte, birçok araştırmacı, hastalığın beynin substantia nigra bölgesinde nöronların dejenerasyonuna yol açan genetik ve çevresel faktörler arasındaki etkileşim sonucu olduğuna inanmaktadır. Sinirlerde meydana gelen hasardan dolayı hastalığın kesin bir tedavisi yoktur (Mounika & Rao, 2021). Uygulanan ilaç tedavileri ile hastanın yaşam kalitesini daha iyi bir hale getirmek amaçlanmaktadır.

Alzheimer'den sonra en sık görülen sinirsel hastalık olan PH, hem motor (Titreme, hareketlerde yavaşlama, denge bozuklukları vb.) hem de motor olmayan (Uyku bozuklukları, zihinsel bozukluklar vb.) semptomlara neden olan ve yaşla birlikte belirtilerinin yaygınlığı artan bir hastalıktır (Lu vd., 2020). Daha çok yaşlı nüfusta yaygın olmasına karşın genç bireylerde de görülmektedir.

PH, yaşlanan toplumda nörolojik bir rahatsızlık olarak büyümeye devam etmektedir. Parkinson Hastalığı Vakfı'na göre 1 milyon Amerikalı Parkinson hastası bulunmakta ve her yıl bu rakam yaklaşık 60.000 artmaktadır. 1,2 milyon Avrupalı ise bu hastalıkla yaşamlarını sürdürürken bu rakamın 2030 yılında iki katına çıkması beklenmektedir (Butt vd., 2018).

2.1.1. Hastalığın Nedenleri

PH, beynin substantia nigra adı verilen bölgesinde sinirler arası iletişimi sağlayan dopamin miktarının azalması sonucu oluşmaktadır (Poewe, 2017). Dopamin kaybının nedenleri kesin olarak bilinmese de çevresel ve genetik faktörlerin rol oynadığı ileri sürülmektedir. Hastalığın seviyesinde değişimler herkeste farklı oranlarda olmaktadır. Genç yaşta hastalık seviyesinde hızlı bir artış gözlemlenebilirken ilerleyen yaşlarda hastalık ilaç tedavisi ile kontrol altına alınabilir (Er, 2021). Dopamin kas sistemini de etkileyerek en basit hareketlerin yapılmasını da engelleyebilmektedir. Substantia nigrada meydana gelen dopamin kaybı sonucunda yaygın olarak görülen başlıca belirtiler; akinezi, bradikinezi, hipokinezi, denge bozuklukları, rijidite (katılık) ve dinlenme titremesidir (Choi vd., 2016).



Şekil 2.1. Substantia Nigra (Hemensaglik, 2018).

2.1.2. Hastalığın Belirtileri

PH, belirtileri kişiye özgü olabilmektedir. Bundan dolayı hastalığın teşhisi açısından farklı belirtileri gözlemlemek ve değerlendirmek oldukça önemlidir. Hem motor belirtiler hem de motor olmayan belirtiler hastalığın teşhisinde erken evreden itibaren önemli rol oynamaktadır (Müller, Assmus, Herlofson, Larsen & Tysnes, 2013).

Titreme, hareketlerde yavaşlama, katılık ve donma gibi hareketin sürekliliğini etkileyen bozukluklar PH motor belirtileri olarak adlandırılmaktadır.

2.1.2.1. Titreme

Parkinson hastalarında titreme dinleme anında oluşmaktadır. Titreşim frekansına bakılarak mevcut durumun PH'tan mı yoksa başka bir fizyolojik rahatsızlıktan mı kaynaklandığı belirlenebilir. Parkinsonlu bireylerde titreşim frekansı 4-6 Hz arasında iken, diğer fizyolojik rahatsızlıklarda 5-12 Hz aralığındadır. Hastalığın ilk evresinde genellikle el ve kollarda tek taraflı olarak başlayan titreme ilerleyen evrelerde baş ve bacak bölgelerinde de görülebilmektedir (Zhang vd.,2020).

2.1.2.2. Hareketlerde Yavaşlama

Hareketlerde yavaşlama (bradikinezi), PH'ın klinik teşhisinde önemli olan özelliklerden biri olmasına rağmen erken evrelerde tespiti zordur. Hastalarda el – parmak hareketleri ve yürüyüş esnasında yavaşlamalar meydana gelmektedir (Lee vd., 2016). Parkinson hastalarında bradikinezi, düşük hız, küçük genlik, düzensiz ritim, kısa duraksamalar ve aşamalı düşüşler içeren karmaşık bir harekettir (Alty vd., 2016).

2.1.2.3. Katılık

Katılık, diğer motor belirtilerden daha hızlı bir şekilde ilerlemektedir. Eklemin hareketine karşılık artan direnç olarak tanımlanmaktadır (Xia, Muthumani, Mao & Powell, 2016). Diz, dirsek, el ve ayak bileği gibi eklem bölgelerinde sık rastlanılan bir belirtidir (Özekmekçi, Apaydın, Oğuz & Zileli, 2013). İlk evrelerde tek taraflı seyreden katılık, ilerleyen evrelerde diğer tarafı da etkileyebilir.

2.1.2.4. Donma

Hareket anında meydana gelen donma Parkinson hastası bireyin hareketini akıcı bir şekilde devam ettirememesi durumudur. Hasta hareketi devam ettirmeye çalışsa da kaslarda meydana gelen donmalar hareketi engellemektedir.

2.1.2.5. Ses ve Konuşma Bozuklukları

Yüz, ağız, gırtlak ve yutak kasları, kas kontrol mekanizmalarında meydana gelen bozulmalardan etkilenmektedir (Gomez vd., 2019; Kurt, Ulukaya & Erdem, 2018; Solana-Lavalle vd., 2021). Yapılan çalışmalarda, ses verileri kolay toplanabilir olduğundan %90 oranında ses ve konuşma bozukluklarına odaklanılmıştır (Ashour vd., 2020; Asmae vd., 2020; Liu vd., 2021; Vásquez-Correa vd., 2018).

2.1.2.6. Yürüyüş ve Denge Bozuklukları

PH'ın ilk belirtilerinden biri olan hareketlerde yavaşlamalar, yürüyüşü de etkilemektedir. Ayrıca ani dönüşlerde düşme ve donmalar oluşmakta olup ilerleyen evrelerde adımlarda küçülme, ayak sürüme ve yürüyüş esnasında eğik bir duruşa sahip olma gibi belirtiler gözlemlenmektedir (Combs vd., 2014).

PH'ta meydana gelen motor olmayan semptomlardan bazıları ise şu şekilde sıralanabilir; ruhsal ve davranışsal bozukluklar, uyku bozuklukları, kabızlık, terleme ve cilt problemleri, ağızda salya birikmesi, düşük tansiyon, ayaklarda şişme, unutkanlık (demans) ve koku alma bozuklukları (Özekmekçi, Apaydın, Oğuz & Zileli, 2013).

2.1.3. Parkinson Hastalığı Tespitinde Kullanılan Yöntemler

Klinik değerlendirmeler sırasında doktorlar tarafından PH'ı tespit etmek amacıyla ölçekler oluşturulmuştur. Gözleme dayalı yapılan testler sonucunda uzman klinisyen hastalığın varlığını teşhis eder veya devam eden hastalık sürecinde şiddetini belirleyebilmektedir.

2.1.3.1.Birleşik Parkinson Hastalığı Değerlendirme Ölçeği (BPHDÖ)

Birleşik Parkinson Hastalığı Değerlendirme Ölçeği (Unified Parkinson's Disease Rating Scale (UPDRS)) hastaların işlevsel yeteneklerine verilen puanlarla ölçülüp 0 ile 4 arasında bir skorlama yapılarak hastalığın derecesinin belirlenmesini sağlayan klinik bir değerlendirme ölçümüdür. En sık kullanılan ve en geçerli yöntemdir. BDHDÖ, nicel gözlemlere değil, klinisyenin gözlemlerine ve hasta ile hasta yakınının o anki beyanına dayanmaktadır (Arroyo-Gallego vd., 2017). Daha sonradan BPHDÖ, hafif dereceli Parkinson hastalarını tespit etmek ve hastalığın yüksek doğrulukla teşhis edilmesi amacıyla geliştirilerek Hareket Bozuklukları Derneği - Birleşik Parkinson Hastalığı Değerlendirme Ölçeği (Movement Disorder Society - Unified Parkinson's Disease Rating Scale (MDS-UPDRS)) adını almıştır.

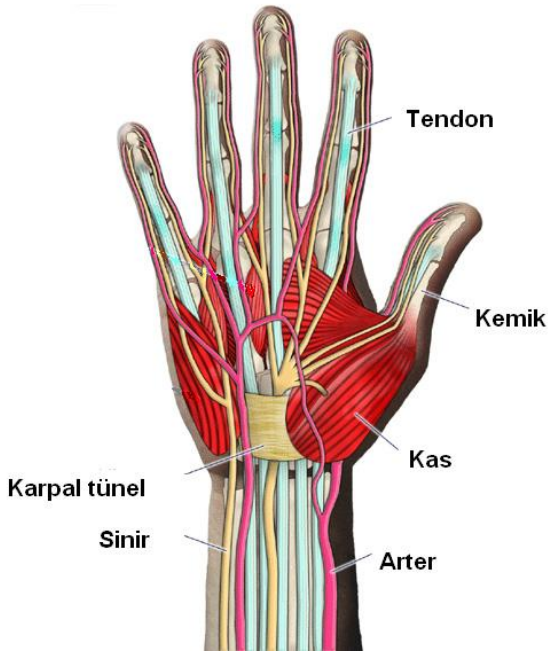
2.1.3.2. Hoehn & Yahr Ölçeği (H&Y)

Hoehn & Yahr Ölçeği, BPHDÖ'den sonra en yaygın kullanılan klinik değerlendirme ölçeğidir. Ancak basitliği ve detay eksikliği nedeniyle kapsamlı bir ölçek olmamakla birlikte motor dışı belirtilerle ilgili bilgi içermemektedir. Bundan dolayı her

ne kadar uygulanabilirliği kolay olsa da çok fazla tercih edilmemektedir (Bhidayasiri & Tarsy 2012).

2.2. El Hareketleri

İnsan eli, kemik ve kaslardan meydana gelen, bireyin yaşamsal faaliyetlerinin yapmasını sağlayan fizyolojik bir yapıdır. Temel yapılarda (kemik, kas ve deri) doğuştan var olan veya çevresel faktörlerle oluşan kişiden kişiye farklılıklar mevcuttur. Bu farklılıklar kişinin yazma (kalemle veya klavyede), imza atma gibi dinamik hareketlerinin kişiye özgü olmasını sağlamaktadır (Li, Zhang & Jin, 2017).



Şekil 2.2. Basit El Anatomisi (Elcerrahi, 2013).

Destek ve hareket sistemi eklemler, kemik doku, kas doku ve kıkırdak dokudan meydana gelmektedir. İnsan vücudu, toplamda 206 kemik ve 3 farklı kas dokudan oluşmakta olup kemik dokunun dörtte biri ellerde bulunmaktadır. 27 kemik ve 30'dan fazla kas ile eller vücudun en karmaşık bölgelerinden biridir (Jones & Lederman, 2006).

Dopamin, sinirler arası mesaj iletimi yapan temel bileşenlerden biridir. Merkezi sinir sisteminden gönderilen mesajlar dopamin yardımı ve sinirler aracılığıyla ellere ulaşır. Kas dokuya ulaşan mesajlar, kas dokunun kasılıp gevşemesi ve kemik dokunun hareketiyle son bulur.

2.3. El Hareketlerinde Bozulmalar ve Parkinson Hastalığı

Beynin substantia nigra bölgesinde dopaminin azalması ile sinirler arası iletişim kötü etkilenmektedir. İletimi geç ya da gerçekleşmeyen sinirsel mesajlar el hareketlerinde düzensizliklere, yavaşlamalara ve sapmalara neden olmaktadır.

2.3.1. Klavye Tuş Vuruş Dinamikleri

Yıllar önce icat edilen telgrafta bireyin cihaza yaptığı vuruşlar ile iletişim sağlanmaktaydı. Ancak bir süre sonra fark edildi ki yapılan tuş vuruşları sadece harfleri değil kişiye özgü yazma davranışları da belirleyebilmektedir. Teknolojinin gelişmesiyle birlikte bu davranışlar bilgisayar klavye tuş vuruşlarının da incelenmesine ortam hazırlamıştır. Kişi tanımlama ve hastalık tespiti gibi alanlarda kullanılan tuş vuruş dinamikleri, tuşa basma ve bırakma arasında gerçekleşen milisaniye cinsiden hesaplanan sürelerdir (Giancardo, Sánchez-Ferro, Butterworth, Mendoza & Hooker, 2015).

Donma süresi, bir tuşa basma ve bırakma arasında geçen süre; gecikme süresi, ard arda basılan iki tuşa basma arasında geçen süre; uçuş süresi ise ard arda basılan iki tuştan ilk tuşu bırakıp ikinci tuşa basma arasında geçen süre olmakla birlikte PH'da meydana gelen yavaşlama, donma ve katılık gibi semptomların belirlenmesinde büyük önem teşkil etmektedir.

2.3.2. Parmak Vuruşu Hareketi (Finger Tapping)

İşaret ve baş parmağın olabildiğince hızlı ve maksimum mesafede birbirine vurulmasıyla gerçekleştirilen parmak vuruş hareketi, PH'ın belirlenmesinde yaygın olarak kullanılan yöntemlerden biridir. Hastalarda dopamin kaybına bağlı olarak artan hareket yavaşlığı ve kaslardaki kasılmalardan dolayı meydana gelen donma ve katılık parmak vuruşlarının hızında düşüşe, parmaklar arasındaki mesafede azalmaya ve vuruşlar sırasında parmak uç noktalarında sapmalara neden olmaktadır (Kim vd., 2011).

2.3.3. El Açma-Kapama Hareketi (Opening-Closing)

El açma-kapama hareketi, Parkinson hastasının dinlenme halindeyken elini belirli bir süre içerisinde en hızlı şekilde maksimum mesafede açıp kapamasıyla gerçekleştirilir. Kasların kontrol mekanizmasında meydana gelen bozulmalardan dolayı el açma-kapama

hızlarında düşüş, belirli bir süre içerisinde el açma-kapama sayısında azalma meydana gelmektedir (Butt vd., 2018).

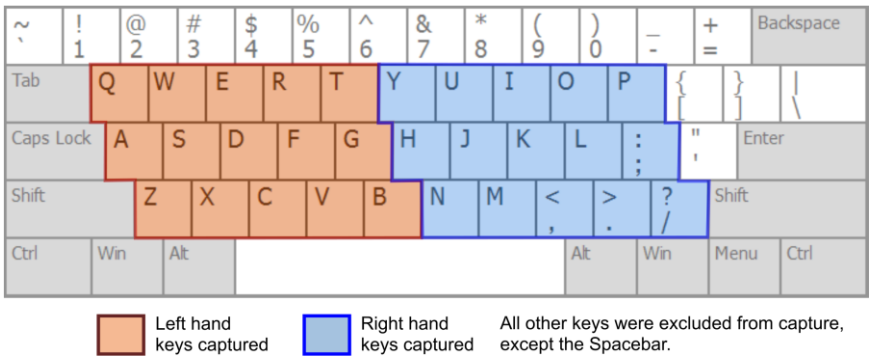
2.3.4. El Döndürme Hareketi (Pronation-Supination)

Dinlenme halindeyken hastanın elini ileri doğru uzatıp en hızlı şekilde, avuç içi yukarı ve aşağı döndürülerek el hızının ve avuç içi merkez noktasının açısının hesaplanmasıdır. Parkinson hastalarında hareketlerde yavaşlama sonucu belirtilen görevleri sağlıklı bireylerden farklı olarak daha düşük hızda gerçekleştirmektedir. Bu nedenle el döndürme hareketi Parkinson hastası - sağlıklı katılımcı ayrımı yapabilmek için iyi bir görevdir (Moshkova vd., 2020).

2.4. El Hareketleri ile Yapılan Parkinson Hastalığı Çalışmaları

El hareketleri, PH teşhisinde en önemli belirteçlerden biridir. Tanı gözleme dayalı olduğundan hastalığın tespiti kullanılan hassas ve nicel teknolojik cihazlar ile daha yüksek doğruluklara ulaşmıştır.

Adams, 2017 yılında yapmış olduğu çalışmasında, klavye tuş verisini ‘Tappy’ adını verdikleri program ile toplayarak katılımcının günlük hayatta klavye yazma davranışlarını incelemiştir (Adams, 2017). Program katılımcılar tarafından bilgisayara indirilerek kurulmuş ve arka planda sürekli çalışarak tuşlara vuruş sürelerini milisaniye cinsinden kaydetmiştir. Ancak katılımcının sol, sağ ya da boşluk tuşuna bastığı süreler bilinirken tam olarak hangi tuşa bastığı bilinmemektedir. Harfler dışında basılan sayı, işaret ve klavyede bulunan diğer tuşlara yapılan vuruş verileri kaydedilmemiştir. Şekil 2.3’te klavye tuş vuruşlarının tarafları gösterilmiştir.



Şekil 2.3. Klavye tuş vuruşlarının tarafları (Turuncu alan sol tuşları; mavi alan sağ tuşları temsil etmektedir) (Adams, 2018).

Tuş vuruş süreleri o anda basılan tuşta yaşanan donma süresi, tuşlar arası geçişte gerçekleşen gecikme ve uçuş süreleri olmak üzere kayıt altına alınmış ancak uçuş süresi çalışmaya dahil edilmemiştir. Tuş geçişleri sol, sağ veya boşluk olarak 3 farklı grupta kayıt altına alınırken bu tuşlar arasında elde edilen geçiş kombinasyonları aşağıdaki tabloda gösterilmiştir.

Çizelge 2.1. Klavye tuş vuruş yönleri.

Sol tuştan sağ tuşa (LR)	Sağ tuştan sağ tuşa (RR)	Boşluk tuşundan sağ tuşa (SR)
Sol tuştan sol tuşa (LL)	Sağ tuştan sol tuşa (RL)	Boşluk tuşundan sol tuşa (SL)
Sol tuştan boşluk tuşuna (LS)	Sağ tuştan boşluk tuşuna (RS)	Boşluk tuşundan boşluk tuşuna (SS)

Çalışmada, belirtilen tuş vuruş geçişlerinin sürelerinden ortalama, standart sapma, basıklık ve çarpıklık özniteliklerini çıkarmıştır. Ancak boşluk tuşu ile yapılan vuruşların hiçbiri çalışmaya dahil edilmemiştir. Çıkartılan öznitelikleri öncelikle donma ve gecikme sürelerine göre ikiye ayırmıştır. Normalizasyon yöntemleri ve lineer diskriminat analizi uyguladıktan sonra 8 farklı sınıflandırma algoritması kullanmışlar ve "ortalama olasılıklar sınıflandırıcısı" adı verdikleri yöntem ile sınıflandırma işlemini gerçekleştirmişlerdir. 217 katılımcı ile gerçekleştirilen çalışmada, hafif derecede, Levodopa ilacı kullanmayan Parkinson hastaları ile sağlıklı bireylerin günlük en az 2000 tuş vuruşu içeren verileri kullanılarak hasta-sağlıklı ayrımı yapılmıştır. Yapılan filtreleme ile 20 PH ve 33 sağlıklı bireyden oluşan alt veri kümesi oluşturulmuştur. Grup A olarak adlandırılan alt veri kümesine ek olarak bağımsız bir veri kümesi olan 'neuroQWERTY MIT-CSXPD' veri seti içerisinden hafif derecede, Levodopa ilacı kullanmayan Parkinson hastaları ile sağlıklı bireylerin günlük en az 500 tuş vuruşu içeren katılımcılar birleştirilmiştir (Giancardo vd., 2016; Goldberger vd., 2000).

A grubu eğitim ve test kümeleri kendi içerisinden seçilirken, B grubu (12 PH, 38 Sağlıklı) eğitim olarak A grubu verilerini, test olarak ise bağımsız verileri kullanılmıştır.

A ve B gruplarına ayrı ayrı uygulanan normalizasyon ve sınıflandırma işlemlerinin ardından A grubunda %100 doğruluk, %100 özgüllük ve %100 hassaslık; B grubunda %94 doğruluk, %95 özgüllük ve %92 hassaslık sonuçları elde edilmiştir. PH'ta hareket yavaşlaması tuş vuruş sürelerinde meydana gelen değişimler ile hesaplanabilmiştir.

Adams 2018 yılında yapmış olduğu bir diğer çalışmada, 'Tappy' programı ile toplanan verilerden her bir el için 50 veya daha fazla devamlı karakterin (yaklaşık 15 kelime) tuş uçuş sürelerini incelemiştir (Adams, 2018). 27 PH ve 49 sağlıklı katılımcı ile yapılan çalışmada, Sinüzoid İnterpolasyon yöntemi ile her bir uçuş süresi arasındaki fark kullanılarak titreme frekansının varlığı tespit edilmeye çalışılmıştır. Parkinson hastaları ve sağlıklı katılımcılar arasında %79 doğruluk, %82 özgüllük ve %73 hassaslık ile titreşim tespit edilirken; sadece hafif derece Parkinson hastalarının dahil olduğu grupta %89 doğruluk, %100 özgüllük ve %78 hassaslık ile titreşim tespit edilmiştir.

Schaap, 'Tappy' klavye tuş vuruş verisini kullanarak tuş vuruş sayısının ve makine öğrenimi yöntemlerinin PH'nın şiddetinin tespitindeki önemini araştırmıştır (Schaap, 2018). Çift boşluk tuşu vuruşları çalışmaya dahil edilmemiştir. Lojistik Regresyon, k-EYK ve Rastgele Orman algoritmalarını veri kümesinde en az 1000, 2000 ve 5000 olacak şekilde incelemiştir. Rastgele Orman ve k-EYK algoritmaları 5000 tuş vuruşu ile kendi sınıflandırma algoritmalarında 1000 ve 2000 tuş vuruşuna göre daha iyi performans göstermiştir. 1000 ve 2000 tuş vuruşu arasında anlamlı bir farklılık ve artan tuş vuruşu ile sınıflandırma performanslarında kademeli bir artış gözlemleyememiştir.

'Tappy' klavye tuş vuruş setini kullanarak araştıran bir diğer çalışma ise Peachap ve arkadaşları (Peachap, Tchiotsop, Louis-Dorr & Wolf, 2020) tarafından yapılmıştır. 217 katılıcımdan oluşan veri seti içerisinden hafif PH olan 22 kişi ve 30 sağlıklı birey ile yeni bir veri kümesi oluşturulmuştur. Bazı katılımcıların günlük vuruş sayıları çok fazla olduğundan büyük veri içeren dosyalar 8000 satırlık parçalara ayrılmıştır. Bunun sonucunda, 286 PH ve 203 sağlıklı tuş vuruş örneklemi elde edilmiştir. Dalgacık dönüşümü yöntemi (Daubechies, ayrık Laguerre ve çift ortoganal) kullanılarak öznitelik çıkarma işlemi gerçekleştirilmiştir. Yapay sinir ağları, DVM, k-EYK ve Lojistik Regresyon sınıflandırma algoritmaları kullanılarak başarım performansları %93,5-100 arasında elde edilmiştir.

Giancardo ve arkadaşları klavye tuş vuruşları veri setinde 42 erken evre Parkinson hastası ve 43 sağlıklı katılımcı ile donma sürelerini araştırmışlardır (Giancardo vd., 2016). Donma veri seti zaman serilerinde meydana gelen değişimler neuro-QWERT adını verdikleri indeks ile gruplandırılarak, ayrık değerleri tahmin etmek için kullanılan Destek Vektör Regresyonu algoritması uygulanmıştır. Algoritma AUC=0,81 performansıyla Parkinson hastalarını sağlıklı katılımcılardan ayırmayı başarmıştır.

Tamson, 2018 yılında yapmış olduğu çalışmasında 'Tappy' klavye tuş vuruş verisini kullanarak tuş vuruş geçişleri arasındaki süreleri istatiksel analiz yöntemleriyle incelemiştir. Veri setlerinin boyutlarının dengesizliğinden dolayı uzun olan dosyalar 1000 tuş vuruşu içerecek şekilde parçalara ayrılmış ve 1000 tuş vuruşundan az olan dosyalar da çıkartılmıştır. Çalışmaya sağdan sola ve soldan sağa yapılan tuş vuruşları dahil edilmiştir. ANOVA (Analysis Of Variance – Varyans Analizi) ve Rastgele Orman yöntemleri ile belirleyici öznitelikler çıkartılmış ve DVM ile k-EYK algoritmaları kullanılarak başarı performansları elde edilmiştir. DVM sonuçları %77,06-79,13 iken, k-EYK sonuçları %76,14-86,64 olarak bulunmuştur (Tamson, 2018).

Alty ve arkadaşları yapmış olduğu çalışmada işaret ve baş parmaklarına yerleştirilen elektromanyetik cihazlar kullanılarak parmakların 30 saniye boyunca birbirine vurulması ile elde eden veriler kullanmışlardır. 49 PH ve 41 sağlıklı katılımcının yer aldığı çalışmada gelişimsel algoritmalar kullanılmış ve %93 sınıflandırma başarısı elde etmişlerdir (Alty vd., 2016).

Gao ve arkadaşları 107 Parkinson hastası, 41 Esansiyel titreme hastası ve 49 sağlıklı katılımcı ile yapmış oldukları bir çalışmada, katılımcılardan 30 saniye boyunca işaret ve baş parmaklarını, parmaklar arası mesafe maksimum olacak şekilde olabildiğince hızlı vurmalarını istemişlerdir. Katılımcıların PD-monitor adı verdikleri skorları parmaklara yerleştirilen elektromanyetik sensörler yardımıyla dijital ortama aktarılmış ve gelişimsel algoritmalar kullanılarak sağ eli baskın Parkinson hastalarında yüksek doğrulukla (≥89,7%) hastalık şiddeti tespit etmişlerdir. Parkinson hastalarını, Esansiyel titremesi olan ve hareket yavaşlaması bulunmayan katılımcılardan ayırt etmeyi başarırken, sağlıklı katılımcılardan ayırt etmede AUC=0,899 değerini elde etmişlerdir (Gao vd., 2018). Literatür çalışmalarına ait özet bilgiler Çizelge 2.2'de verilmiştir.

Çizelge 2.2. Literatür çalışmaları.

Çalışmanın Adı	Kişi Sayısı	Uygulanan Yöntemler	Kullanılan Cihaz	Sonuç
Adams (2017)	32 PH 61 Sağlıklı	8 Farklı Sınıflandırma Algoritması	Klavye	%94-100 doğruluk
Adams (2018)	27 PH 49 Sağlıklı	Sinüzoid İnterpolasyon	Klavye	%79-89 doğruluk
Schaap (2018)	109 PH 35 Sağlıklı	Lojistik Regresyon, Rastgele Orman, k-EYK	Klavye	RO ve k-EYK algoritmalarında 5000 tuş vuruşu, diğer tuş vuruşlarına göre en iyi performansı göstermiştir.
Peachap, Tchiotsop, Louis-Dorr & Wolf (2020).	22 PH 30 Sağlıklı	Yapay sinir ağları, DVM, k-EYK Lojistik Regresyon	Klavye	%93,5-100 doğruluk
Giancardo, Sánchez-Ferro, Butterworth, Mendoza & Hooker (2015).	42 PH 43 Sağlıklı	Destek Vektör Regresyonu	Klavye	AUC = 0,81
Tamson (2018)	147 PH 49 Sağlıklı	DVM k-EYK	Klavye	DVM, %77.06-79.13 doğruluk k-EYK, %76,14-86.64 doğruluk
Alty vd. (2016)	49 PH 41 Sağlıklı	Gelişimsel Algoritmalar	Elektro manyetik sensör	%93 doğruluk
Gao vd. (2018)	107 PH 41 Esansiyel Titreme 49 Sağlıklı	Gelişimsel Algoritmalar	Elektro manyetik sensör	Hastalık Şiddeti Doğruluk ≥89,7% Sağlıklı Katılımcı AUC=0,899

BÖLÜM 3

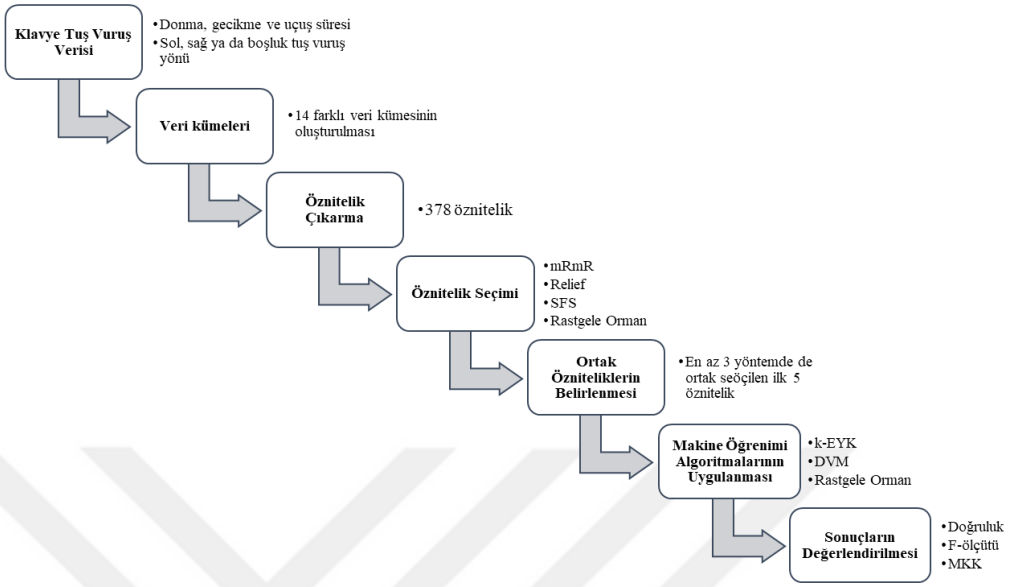
MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK PARMAK VURUŞ HAREKETİ VERİLERİ İLE PARKİNSON HASTALIĞI TESPİTİ

3.1. Makine Öğrenimi

Yapay zekanın bir alt dalı olan Makine Öğrenimi, verilerden anlamlı bilgiler edinebilmek amacıyla matematiksel modellerin oluşturulmasıdır. Algoritmalar ile model, veriye göre değişkenlik gösteren parametreler yardımıyla var olan veriye uygun hale getirilerek 'öğrenme' işlemi gerçekleştirilir. Sonraki aşamada modelin daha önce tanımadığı veriler modele işlenerek verilerin tahmin ve yorumlanması yapılmaktadır (VanderPlas, 2016).

Makine öğrenimi modeli oluşturulmadan önce verinin ön işlemeden geçmesi gerekmektedir. Veri analizi sonuçlarının kalitesini artırmak amacıyla veri integrasyonu, veri temizleme, veri dönüştürme ve veri indirgeme gibi ön işlemler yapılmaktadır (Akpınar, 2017). Veri kullanılabilir hale getirildikten sonra nicel ya da görsel olan veriden, her türlü bilgiyi içeren öznitelikler (minimum, maksimum, ortalama vb.) algoritmalar ile çıkartılır. Farklı veri kümelerinde farklı öznitelikler modelin performansını etkileyebildiğinden verinin yapısı incelenerek en başarılı özniteliklerin belirlenmesi çok önemlidir. Çıkartılan öznitelikler modelin veriyi daha iyi anlayabilmesi için çeşitli normalizasyon yöntemlerinden geçirilerek belirli bir aralığa sıkıştırılır. Uygun sınıflandırma algoritmaları kullanılarak öznitelikler yardımıyla çıkarım yapılır. Model, hastalık derecesinin belirlenmesi veya hasta-sağlıklı ayrımı yapılması gibi farklı

yapılarda olabilir. Bu çalışmada uygulanacak olan makine öğrenmesi basamakları Şekil 3.1'de gösterilmiştir.



Şekil 3.1. Genel akış şeması.

3.2 Verilerin Elde Edilmesi

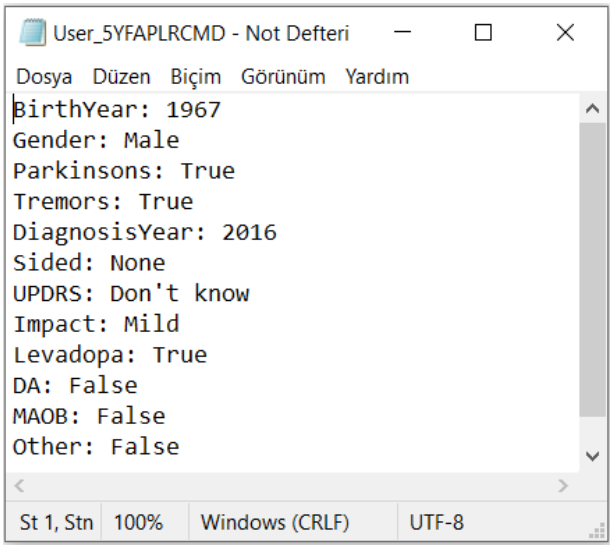
Literatür çalışmalarında 'Tappy' adlı programla toplanan klavye tuş vuruş verileri Physionet veri tabanından elde edilmiştir (Goldberger vd., 2000). 162 Parkinson hastasını ve sağlıklı 55 katılımcıyı içeren veri kümesinde klavye tuş vuruşlarına ait donma, gecikme ve uçuş süreleri yer almaktadır. Parkinson hastası 2 kişiye ait veri dosyaları bulunmadığından veri kümesine dahil edilmemiştir.

Adams'ın 2017 yılında yapmış olduğu çalışmada tuş vuruş donma ve gecikme sürelerine ait 27 öznitelik incelenmiş ancak uçuş süresine ve boşluk tuşuna basma süresine ait öznitelikler dahil edilmemiştir. Çizelge 3.1'de veri seti özeti verilmiştir.

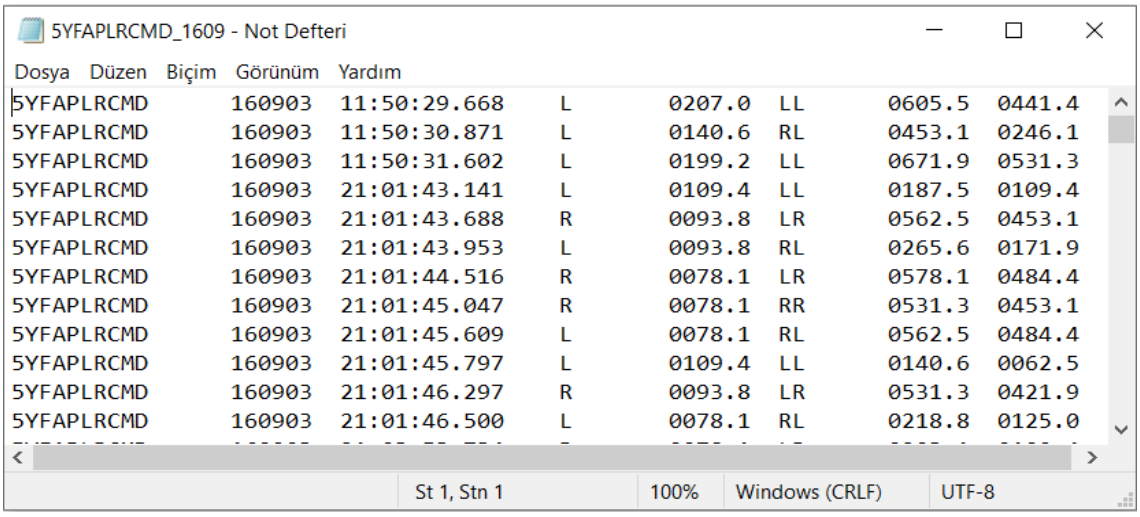
Çizelge 3.1. Veri seti özellikleri.

	Parkinson	**Sağlıklı**
Katılımcı Sayısı	160	55
Cinsiyet		
Kadın	77	24
Erkek	83	31
Yaş		
Ortalama	68.52	68.95
Aralık	32 - 90	31 - 89
Standart Sapma	10.02	11.75
Bilgi yok	31 kişi	6 kişi
Hastalık Süresi (Yıl)	0 - 26	-
Hastalık Şiddeti		
Hafif	64	-
Orta	68	-
Şiddetli	24	-
Bilgi yok	4	-
Levodopa		
Var	106	-
Yok	54	-
Örneklem Sayısı	425	129

Veri setinde her katılımcı için iki ayrı dosya bulunmakta olup, bu dosyalar rastgele 10 karakterden oluşan kişi isimleri adı altında oluşturulmuştur. İlk dosya katılımcının sağlık bilgilerini (Doğum yılı, cinsiyet, titreme (var/yok), hastalık yılı, etkilenen taraf (sağ/sol), BPHDÖ skoru, hastalık şiddeti ve ilaç kullanım durumları), ikinci dosya ise tuş vuruş sürelerini ve tuş vuruş yönlerini içermektedir. Örnek sağlık bilgisi ve tuş vuruş verisi dosyaları Şekil 3.2 ve Şekil 3.3'te verilmiştir.



Şekil 3.2. Katılımcı Sağlık Bilgileri.



Şekil 3.3. Katılımcı Tuş Vuruş Verileri.

Veri setinde her katılımcının birden fazla güne ait verisi bulunmaktaydı. Ayrıca günlük tuş vuruş sayıları kişiden kişiye farklılık göstermekteydi. Veri setinde meydana gelen bu dengesizliğin önüne geçmek amacıyla bu çalışmada, her katılımcının en az 2000 tuş vuruşu içeren bir günlük verisi ele alınarak incelenmiştir. Yapılan filtreleme işlemiyle 88 Parkinson hastası ve 30 sağlıklı katılımcıdan oluşan yeni bir veri kümesi oluşturulmuştur.

Yapılan çalışmaların genelinde hafif Parkinson hastaları dahil edilirken bu çalışmada katılımcıların sağlık bilgileri dikkate alınarak 16 farklı veri kümesinde sınıflandırma performansları incelenmiştir.

3.3. Öznitelik Çıkarımı

PH'ta dopaminin dejenere olmasından dolayı sinirsel iletimde gecikmeler ve kas kontrol mekanizmasında bozulmalar meydana gelmektedir. PH, ilk evrelerde el ve kollar gibi üst uzuvlarda titreme, hareketlerde yavaşlama, katılık ve donma ile karakterizedir. Klavye tuş vuruş verisi donma, gecikme ve uçuş sürelerini içermektedir.

Veri seti içerisinde hangi tuşa basıldığı ve bir önceki bırakılan tuş arasındaki bilgiler (LL, LR, LS, RL, RR, RS, SL, SR, SS) de mevcuttur.

Yapılan çalışmaların çoğunda boşluk tuşuna ait vuruşlar veri setine dahil edilmemiştir. Ayrıca yapılan vuruşların hangi taraftaki tuşlara yapıldığı bilinirken, sağ ya da sol elle mi yapıldığı hakkında bilgi mevcut değildir. Boşluk tuş vuruşu klavyede yazma

davranışında genellikle baş parmak ve nadiren işaret parmak ile gerçekleştirilmektedir. Parmak vuruşu testi ile yapılan çalışmalar göstermiştir ki işaret ve baş parmakta meydana gelen hareket yavaşlaması, titreme ve donma gibi belirtiler sağlıklı ve Parkinson hastalarını yüksek başarım ile ayırmıştır. Bu nedenle bu çalışmada boşluk tuşuna ait veriler de dahil edilerek hasta – sağlıklı ayırma işlemindeki önemi araştırılmıştır. Donma, katılık ve uçuş süreleri vuruş yönlerine ayrılarak aşağıda verilen öznitelikler her yön için ayrı ayrı hesaplanmıştır.

Çizelge 3.2. Çıkarılan öznitelikler.

Vuruş Süresi	Vuruş Yönü	ÖZNİTELİKLER	Katılımcı 1	…	Katılımcı n
Donma - Gecikme - Uçuş	LL LR LS RL RR RS SL SR SS	1. Minimum			
		2. Maksimum			
		3. Aritmetik Ortalama			
		4. Medyan			
		5. Standart Sapma			
		6. Aralık			
		7. Entropi			
		8. Basıklık			
		9. Çarpıklık			
		10. Kolmogorov Karmaşıklığı			
		11. Lempel-Ziv Karmaşıklığı			
		12. Örnek Entropi			
		13. Higuchi Fraktal Boyut Analizi			
		14. Katz Fraktal Boyut Analizi			

3.3.1. Minimum

Verinin içerisindeki en küçük sayısal değeri temsil etmektedir. Uzun veri setlerinde alt sınırı belirlemek için kullanılmaktadır.

3.3.2. Maksimum

Veri içerisinde var olan sayısal değerlerden en büyüğüdür. Üst sınırın belirlenmesiyle veri modellemede sık kullanılan özniteliklerden biridir.

3.3.3. Ortalama

Verinin sayısal değerlerinin toplamının veri sayısına bölünmesiyle elde edilir. Veri sayısının n olduğu bir veri kümesinde ortalama Denklem 3.1. şeklinde hesaplanır ve istatiksel analiz yöntemlerinde kullanılır.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3.1}$$

3.3.4. Medyan

Merkezi konum ölçüsü olarak tanımlanan medyan, örneklem serisinin küçükten büyüğe sıralanarak veri setini ortadan ikiye bölen değerdir. n tane veri içerisinde n tek sayı ise medyan olarak ortadaki değer alınırken, n çift ise ortadaki iki değerin aritmetik ortalaması alınarak hesaplanır.

3.3.5. Standart Sapma

Veri setinde alt ve üst sınırlarda yığılmalara neden olacak şekilde değerler bulunabilmektedir. Aritmetik ortalamayı etkileyen bu değerlerin birbirine yakınlık ve uzaklıkları arasındaki bağlantıyı hesaplayan standart sapma, ortalamadan sapma miktarı ile risk miktarını da belirleyebilir. Ortalaması $\bar{x}$ olan n uzunlukta bir veri kümesinde standart sapma şu şekilde hesaplanmaktadır.

$$s = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{1/2} \tag{3.2}$$

3.3.6. Aralık

Veri setinin minimum ve maksimum değerleri arasındaki farktır. Verinin en küçük ve en büyük değerleri arasındaki değişim, diğer verilerle kıyaslama yapılarak yorumlanır.

$$aralık = x_{maks} - x_{min} \tag{3.3}$$

3.3.7. Entropi

Verinin düzensizliğinin belirlenmesini sağlayan entropi, Claude Elwood Shannon tarafından 1948 yılında yayımlanan "A Mathematical Theory of Communication" adlı çalışmada incelenmiştir (Shannon, 1948). Rastgele seçilen bir değişkendeki düzensizliği temsil eden entropi, 0-1 aralığında değerler alarak eğri ile temsil edilip eğrinin keskinliğine göre yorumlanır. Düzensizliğin artmasıyla elde edilen kesin bilgide artış gözlemlenirken azalmasıyla bilgi kazanımında düşüşler meydana gelmektedir.

3.3.8. Basıklık

İki dağılımı karşılaştırmak için kullanılan ölçümlerden biridir. Normal dağılımda 0 değerini almaktadır. Basıklık katsayısı 3'ten küçük ise basık bir eğri, büyük ise sivri bir eğriden söz edilmektedir.

3.3.9. Çarpıklık

Çarpıklık, veri kümesindeki değerlerin dağılımının ne kadar simetrik olduğunun bilgisini vermektedir. Tam bir simetrinin olduğu normal dağılımda çarpıklık değeri 0'a eşitken çarpıklık katsayısı pozitif ise sağa çarpık; negatif ise sola çarpık bir eğri elde edilir. Çarpıklık değerleri -0,5 ile 0,5 arasında ise simetrik; -1 ile -0,5 ve 0,5 ile 1 arasında değerler alıyorsa düşük çarpıklık; -1'den küçük ve 1'den büyük değerler alıyorsa yüksek çarpıklık bulunmaktadır.

3.3.10. Kolmogorov Karmaşıklığı

Verinin düzensizliğini mümkün olan en kısa şekilde temsil eden programın uzunluğudur. Kodun uzunluğu verinin karmaşıklığına bağlıdır. Veri ne kadar karmaşık bir yapıya sahipse veriyi oluşturan program kodu da o kadar uzundur (Akar, Kara, Agambayev & Bilgiç, 2015).

3.3.11. Lempel-Ziv Karmaşıklığı

Veri kümesinin çeşitliliğini ölçerek verinin ne kadar karmaşık olduğunu hesaplamaktadır. Verinin ortalama değeri alınarak ortalamanın üstündeki değerler 1'e, altındaki değerler ise 0'a eşitlenerek veri seti ikililere dönüştürülür. İkili diziden oluşan veri ilk değerden başlanarak her seferinde önceden belirlenen değerlerle benzerlikler incelenerek bir alt dizi oluşturulur. Karmaşıklık, oluşturulan yeni veri değerlerinin

uzunluğu ile elde edilir. Veri içerisindeki değerler birbirine ne kadar benzerlik gösteriyorsa veri o kadar az karmaşıklığa sahiptir. (Lempel & Ziv, 1976).

3.3.12. Örnek Entropi

Yapılan çalışmalarda, sinyalin karmaşıklığını belirleyebilmek için Yaklaşık entropi (Approximate Entropy) kullanılmıştır. Yaklaşık entropi, kısa veri kümelerinde verinin karmaşıklığı hakkında nicel bilgiler vermesinden dolayı popüler yöntemlerden biridir. Richman ve Moorman tarafından yaklaşık entropinin geliştirilmiş bir versiyonu olan Örnek entropi (Sample Entropy) tanıtılmıştır (Richman & Moorman, 2000). Örnek entropi, yaklaşık entropi gibi kolay uygulanabilir bir yöntemdir. Ayrıca yaklaşık entropiye ait veri kümesinin uzunluğuna ve tutarlılığının yoksunluğuna bağlı sınırlamaları ortadan kaldırmaktadır (Chen, Solomon & Chon, 2006).

3.3.13. Higuchi Fraktal Boyut Analizi

Fraktal boyut analizi, veri seti içerisindeki benzerliği ölçmektedir. Doğrusal olmayan veri setindeki örüntülerin veri içerisindeki tekrarlarının sayısını belirlemeye çalışarak verinin kendisiyle benzerliğini ölçer, bu da verinin ne kadar karmaşık bir yapıya sahip olduğu hakkında bilgi verir. Higuchi Fraktal Boyut Analizi (HFBA), fraktal boyutu zaman serileri üzerinden tanımlayarak veri değerlerinin benzerliklerini inceler (Atasoy, Yıldırım, Yıldırım & Kutlu, 2014).

3.3.14. Katz Fraktal Boyut Analizi

Katz Fraktal Boyut Analizi (KFBA), veri setindeki ardışık değerler arasındaki mesafeyi kullanarak verinin benzerliğini inceler. Daha iyi sonuçlar almak için uzun veri setlerinde kullanılmaktadır (Garner, De Souza & Vanderlei, 2018).

3.4. Normalizasyon Yöntemleri

Veri setinde yer alan değerler bazı durumlarda uç değerlere sahip olmaktadır. Sınıflandırma sırasında kullanılacak olan bu değerler sınıflandırma performansını etkileyebilmektedir. Bu amaçla veri seti belirli bir aralığa sıkıştırılarak dönüştürülür (Akpınar, 2017). Yapılan bu işlem veri içerisindeki düzeni sağlayarak sınıflandırma performansını yükseltir.

3.4.1. Minimum – Maksimum Normalizasyon

Veri kümesindeki değerlerin, sınıflandırmayı daha verimli hale getirmesi amacıyla belirli bir aralığa sıkıştırılmasıyla elde edilir. Bu çalışmada değerler [-1 +1] aralığına sıkıştırılmıştır. En küçük değer -1 ile, en büyük değer 1 ile ve ara değerler bu aralıkta uygun değerler alacak şekilde değiştirilmiştir. Bir x verisine ait yeni değer Denklem 3.4 ile hesaplanmaktadır.

$$x = \frac{x - min_{veri}}{max_{veri} - min_{veri}} (yeni_maks_{veri} - yeni_min_{veri}) + yeni_min_{veri} \quad (3.4)$$

3.4.2. Z-skor Normalizasyon

Normalleştirme yöntemlerinden biri olan z-skor normalleştirme, verinin ortalaması ve standart sapması yardımıyla veriye yeni bir değer atayarak veri setini [-1 +1] veya [0, +1] arasına sıkıştırmaya çalışmaktadır. Ancak hesaplama sonucunda yeni değerler sınır değerleri aşabilmektedir. Ortalamanın $\bar{x}$, standart sapmanın s olduğu bir veride x verisine ait z-skorun hesaplanması Denklem 3.5'te verilmiştir (Akpınar, 2017).

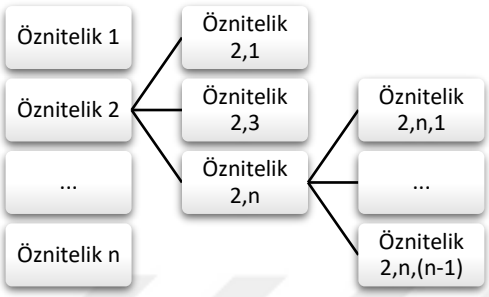
$$x_{yeni} = \frac{x - \bar{x}}{s} \quad (3.5)$$

3.5. Öznitelik Seçim Yöntemleri

Özniteliklerin sayısı gerek veri setinin büyüklüğünden gerekse kullanılan öznitelik çıkarma yöntemlerinin çokluğundan dolayı fazla sayıda olabilir. Elde edilen özniteliklerin her biri sınıflandırıcıya dahil edildiğinde yüksek ayırma gücüne sahip olmayan öznitelikler sınıflandırma performansını etkileyebilmektedir. Bundan dolayı, öznitelik seçim yöntemleri geliştirilerek başarı performansının artırılması amaçlanmıştır. Bu yöntemler, ayırma gücüne göre öznitelikleri sıralayıp, sınıflandırmada kullanmak amacıyla en iyi olanları seçmektedir.

Öznitelik seçme yöntemleri şu şekilde kategorize edilebilir: Filtre Yöntemler, Sarıcı Yöntemler ve Gömülü Yöntemler. Filtre yöntemleri, Ki-Kare Test Yöntemi ve Pearson Korelasyon ölçütünü kullanarak öznitelik ile veri arasındaki bağlantıyı dikkate alıp en iyi olanları filtreleme yapma işlemidir. Ölçeklenebilir ve hızlı bir hesaplamaya sahiptir. Sarıcı yöntemlerde, öncelikle tüm özniteliklere bir başarı puanı atanır. Yüksek başarı puanına sahip öznitelik ile diğer öznitelikler teker teker gruplandırılır. Gruplar içerisinde en iyi performansı olan seçilerek, diğer özniteliklerin teker teker eklenmesiyle

tekrar gruplandırma yapılır ve bu şekilde özniteliklerin hepsi performans gücüne göre sıralanır (Şekil 3.4). Özniteliklerin teker teker değil kombinasyonunun kullanılması açısından avantajlı bir yöntem iken zaman tüketiminin fazla ve ezbere öğrenme tehlikesi açısından dezavantajlı bir yöntemdir. Gömülü yöntemlerde ise en başarılı öznitelikler sınıflandırma sırasında algoritmanın öznitelikleri önemine göre sıralaması ile elde edilir.



Şekil 3.4. Sarıcı öznitelik seçme yöntemi temsili.

3.5.1. Minimum Artıklık Maksimum Bağlantı [Minimum Redundancy Maximum Relevance (mRmR)]

Filtre bazlı öznitelik seçim yöntemlerinden biridir. Ding ve Peng tarafından 2005 yılında önerilen minimum artıklık ve maksimum bağlantı öznitelik seçim yöntemi, sürekli yinelenerek ve her yinelemede seçilen öznitelikler ile minimum artıklık ve ilgili sınıf bilgisiyle maksimum bağlantı prensibine dayalı olarak çalışmaktadır (Ding & Peng, 2005).

Öznitelikler, sınıflandırma gücüne göre önceden puanlanarak maksimum bağlantılı olan öznitelik seçilir. Ardından diğer öznitelikler arasından sınıf bilgisiyle maksimum bağlantıya sahip, önceden seçilmiş olan özniteliklerle minimum fazlalığa sahip olan öznitelik seçilir (Gülgezen, Çataltepe & Yu).

Öznitelikler arasındaki bağlantıyı ölçmek için Spearman ve Pearson korelasyon katsayısı kullanılmaktadır. Bu çalışmada, Pearson korelasyon katsayısı benzerlik ölçeği kullanılmıştır. X ve Y arasındaki benzerlik ölçütü, Pearson korelasyon katsayısına ait Denklem 3.6 ile hesaplanmıştır. n toplam öznitelik sayısı, s_X ve s_Y özniteliklerin standart sapması, $\bar{X}$ ve $\bar{Y}$ özniteliklerin aritmetik ortalamasını temsil etmektedir (Akpınar, 2017).

$$Pearson\ korelasyonu_{XY} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{(n-1)s_X s_Y} \tag{3.6}$$

3.5.2. RELIEF

1992 yılında Kira ve Rendell tarafından geliştirilen RELIEF algoritması filtre bazlı öznitelik seçim yöntemlerinden biridir (Kira & Rendell, 1992). RELIEF algoritması başlangıçta tüm özniteliklere 0 ağırlığını atayarak içlerinden rastgele birini seçmektedir. Seçilen veriyle aynı sınıfa ait, sınıf sayısından bir eksik olacak sayıda veri bulunur, devamında aynı işlem farklı sınıfa ait veriler için gerçekleştirilir. Bu veriler yardımıyla öznitelikler ağırlıklandırılır ve yeni bir öznitelik alt grubu oluşturulur. Seçilmeyen öznitelikler elenir (Kaynar, Arslan, Görmez & Işık, 2018).

3.5.3. Ardışık İleri Yönde Seçim [Sequential Forward Feature Selection (SFS)]

Basitliği nedeniyle sık kullanılan algoritmalardan biri olan Ardışık İleri Yönde Seçim Algoritması, sarıcı öznitelik seçim yöntemlerinden biridir. Boş bir öznitelik kümesi oluşturarak çalışmaya başlar. Gelişmiş fonksiyonlar yardımıyla özniteliklerin ortalama kare hatasını en aza indirip özniteliği kümeye dahil eder. Bu işlem yinelenerek devam eder ve oluşturulan yeni öznitelik kümesinin eklenen her yeni öznitelikle karşılaştırıldığında minimum bir sınıflandırma hatası üretmesine dikkat edilir (Marcano-Cedeño, Quintanilla-Domínguez, Cortina-Januchs & Andina, 2010).

3.5.4. Rastgele Orman

Gömülü öznitelik seçim yöntemlerinden biri olan Rastgele Orman (RO) algoritmasında, öznitelik seçilimi sınıflandırma esnasında gerçekleştirilir. Algoritma belirlenen sayıda ağaç oluşturur. Ardından ağaçların verdiği sonuçlar incelenir ve çoğunluğun oyu ile bir seçilim yapılır. Yapılan bu oylama sayesinde tahminin doğruluğu güçlendirilir.

Hem regresyon hem de sınıflandırma problemlerinde kullanılmaktadır. Farklı ağaç yapılarına ait farklı veriler kullanıldığından dolayı aşırı öğrenmeye karşı güçlüdür.

Rastgele orman yönteminde veri kümesi eğitim ve test kümesi olmak üzere ikiye ayrılır. Eğitim seti, inBaG adlandırılarak veri kümesinin 2/3'ten oluşurken, test seti Out-Of-Bag (OOB) olarak adlandırılır ve veri kümesinin 1/3'den meydana gelmektedir (Kurt, 2019).

Veri kümesinden üretilecek karar ağacı sayısı kadar bootstrap yöntemi kullanılarak örneklem uzayı oluşturulur. Örneklem uzayı kadar test ve eğitim verisi

ayrılarak test verisi ile hata oranları tahmini yapılır ve ortalamaları alınarak test kümesi hata oranı hesaplanır. Öznitelikler karar ağacı yapısına yerleştirilerek uygun sınıflara atanır. Hata oranı yardımıyla ağaçlara verilen ağırlık derecesi özniteliklerin ağırlıklarını belirler. Hata ne kadar yüksekse ağırlık derecesi o kadar düşüktür. Yani özniteliğin önemi düşer. Yapılan oylama ile en önemli öznitelikler en çok oyu alan ağaçtan seçilirler.

3.6. Sınıflandırma Yöntemleri

Sınıflandırma yöntemleri, veri kümesi içerisinde hangi sınıfa ait olduğu bilinmeyen verinin belirleyici özellikleri yardımıyla modellenmesi ve sınıf tayinin yapılmasından oluşmaktadır. Veri setinin yapısı ve yapılacak olan sınıflandırmanın alanına göre algoritmalar seçilerek sınıflandırma performansı arttırılabilir.

Bu çalışmada kullanılan sınıflandırma algoritmaları, Destek Vektör Makineleri, Rastgele Orman ve k-En Yakın Komşudur.

3.6.1. k-En Yakın Komşu Algoritması

Cover ve Hart tarafından 1967 yılında önerilen k-en yakın komşu algoritması (k-EYK), denetimli öğrenme yöntemleri arasında en çok tercih edilen yöntemlerden biridir (Cover & Hart, 1967; Çalışkan & Soğukpınar, 2008). Veri kümesinde yer alan değerlerin birbirine olan uzaklıkları incelenerek, seçilen bir test verisinin belirlenen komşuluk sayısına (k) göre en yakın olan komşulara bakılır. Seçilen komşulukların sınıf bilgileri oy çokluğu ile hangi sınıfa ait ise veri o sınıfa atanır.

Veriler arasındaki mesafeyi belirleyebilmek için uzaklık ölçülerinden yararlanılmaktadır. Bu çalışmada, Euclid, City-Block (Manhattan), Chebyshev ve Minkowski uzaklık ölçüleri kullanılmıştır. Ölçümlere ait denklemler Denklem 3.7, Denklem 3.8, Denklem 3.9 ve Denklem 3.10'da belirtilmiştir. (Akpınar, 2017).

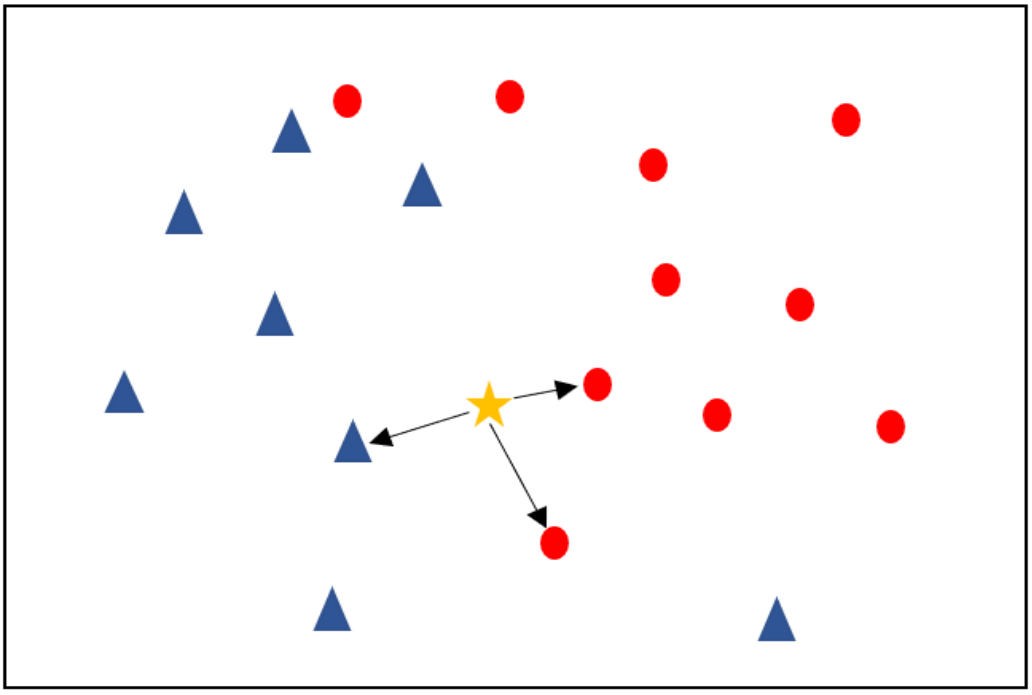
$$Euclid\ Uzaklığı_{\ a,b} = \sqrt{\sum_{i=1}^{n}(x_{ai} - x_{bi})^2} \tag{3.7}$$

$$City - Block\ Uzaklığı_{\ a,b} = \sum_a |x_a - x_b| \tag{3.8}$$

$$Chebyshev\ Uzaklığı_{a,b} = \max_a(|x_a - x_b|) \tag{3.9}$$

$$Minkowski\ Uzaklığı_{a,b} = \sqrt[P]{\sum_{i=1}^{n}|x_{ai} - x_{bi}|^P} \tag{3.10}$$

Test verisinin veriye eklenmesinin ardından veri kümesindeki tüm değerlere olan uzaklıkları hesaplanır. Belirlenen komşuluk değeri ile en yakın komşular bulunur. Komşuluk sayısı eşit oylamanın önüne geçmek için tek sayı olarak seçilmektedir. Sık kullanılan komşuluk (k) değerleri 3,5,7 iken bu çalışmada k değerleri 3,5,7,9 seçilerek, uzaklık ölçümleri her bir komşuluk için hesaplanmış ve sınıflandırma performansları incelenmiştir. Test verisi en yakın komşulara ait sınıf bilgileri içerisinden en çok var olan sınıfa atanarak sınıflandırma işlemi tamamlanır (Şekil 3.5).



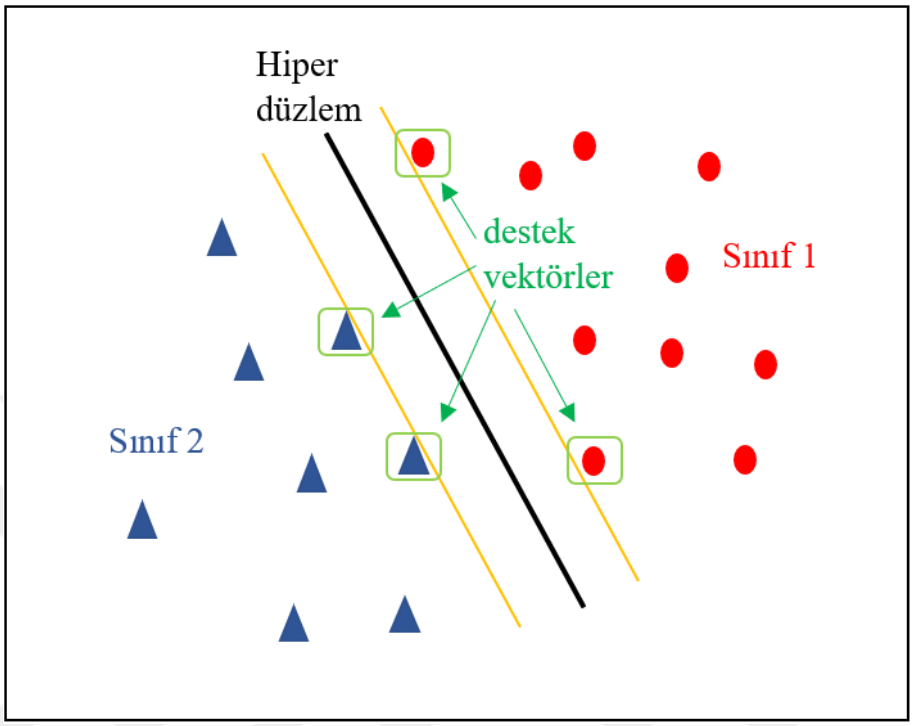
Şekil 3.5. k-En Yakın Komşu Algoritması ile k=3 komşuluğunda verinin sınıflandırılması.

3.6.2. Destek Vektör Makineleri

Denetimli öğrenme yöntemlerinden biri olan Destek Vektör Makineleri (DVM), veriyi en iyi şekilde ayıracak hiper düzlemi bularak sınıflandırmayı amaçlanmaktadır (Pisner & Schnyer, 2020). Doğrusal iki sınıfa sahip veri kümeleri için tasarlanmışken, daha sonradan doğrusal olmayan çok sınıflı veri kümelerinin sınıflandırmalarında kullanılmak üzere geliştirilmiştir.

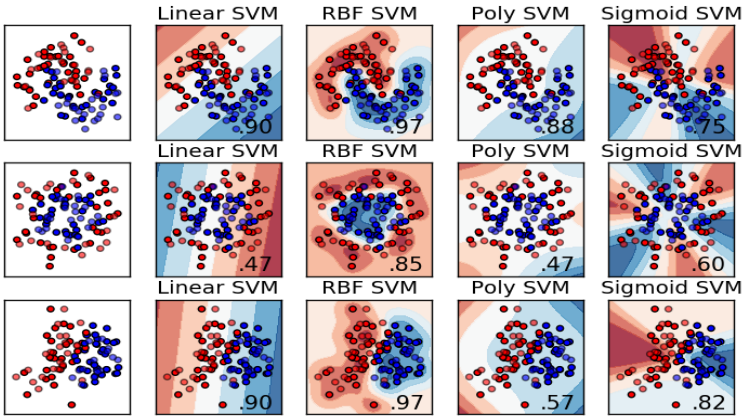
Yapısından kaynaklı nedenlerden dolayı veri kümesi, doğrusal bir şekilde ayrılmayabilir. Bu durumda veri kümesinin boyutu yükseltilerek ayırma işlemi bir düzlem yardımıyla yapılır. Düzlemin en yakınında olan veriler destek vektörleri meydana

getirmektedir. Veri kümesi karmaşıklığı ve üretilen vektörlerin sayısı doğru orantılıdır. Hiper düzlem seçilirken, kendisine en yakın noktalar arasındaki uzaklığı maksimum yapabilmesi beklenmektedir (Şekil 3.6).



Şekil 3.6. DVM hiper düzlem ve destek vektörlerin belirlenmesi.

Veri kümesinin karmaşıklığı bazı durumlarda sınıfların doğrusal bir düzlemle ayrılmasını imkânsız hale getirebilir. Bu durumda, çekirdek fonksiyonları kullanılarak sınıf boyutu arttırılır ve ayırma işlemi yapılır. Çekirdek fonksiyonlarının doğrusal, polinom, radyal (RBF), Laplace, Anova, Sigmoid, Bessel gibi çeşitleri bulunmaktadır. Bu çalışmada, doğrusal, polinom, RBF ve sigmoid çekirdek fonksiyonları kullanılmış ve t parametresi ile adlandırılmıştır (t=0, lineer; t=1, polinom; t=2, RBF; t=3, sigmoid). Şekil 3.7'de 3 farklı veri kümesinde kullanılan çekirdek fonksiyonlarının performansları gösterilmektedir.



Şekil 3.7. DVM çekirdek fonksiyonları (Bilogur, 2017).

3.6.2. Rastgele Orman

Bir öznitelik seçim yöntemi dışında sınıflandırma algoritması olarak da kullanılan Rastgele Orman (RO) algoritması, ağaç tabanlı bir algoritmadır (Breiman, 2001). DVM gibi sınıflandırma algoritmalarından daha keskin sonuçlar vermektedir. Bunun bir nedeni, oldukça büyük ağaç yapıları kurarak ağaçlar arasında düşük korelasyon oluşturmasıdır.

Veri kümesinden rastgele seçilen her veri için bir karar ağacı kurulur. Her karar ağacından elde edilen tahminlerin oy çokluğu ile sınıflandırma yapılır.

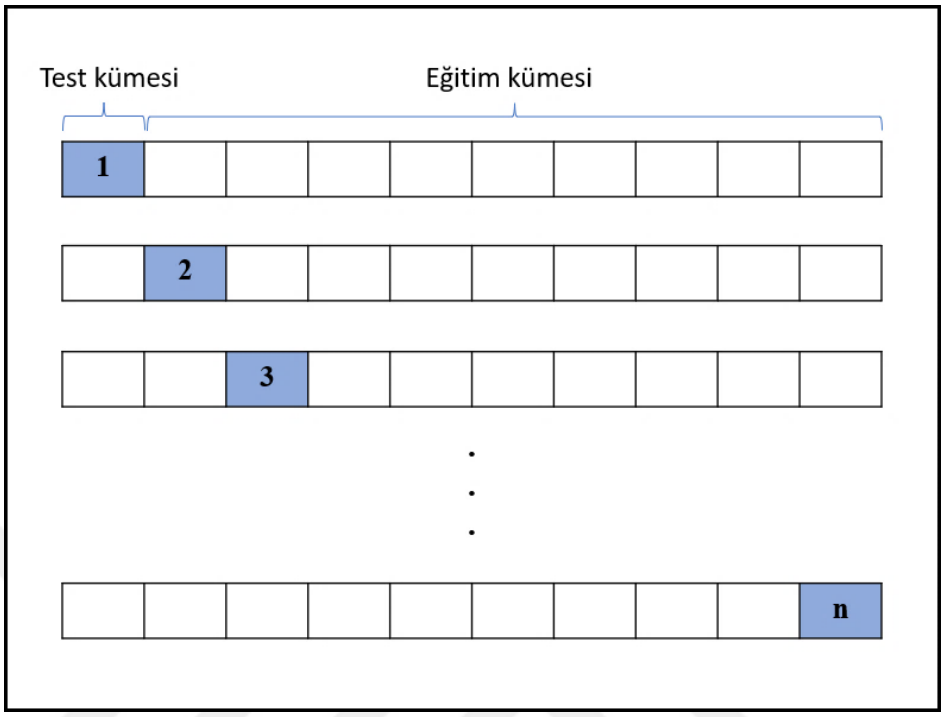
Bu çalışmada Rastgele Orman Algoritması her veri kümesi için 100 defa çalıştırılmış ve çıkan sonuçların ortalaması baz alınarak incelenmiştir.

3.7. Birini Dışarda Bırakma [Leave-one-out (LOO)] Çapraz Doğrulama Yöntemi

Veriyi eğitim ve test kümesi olmak üzere ikiye ayırarak modelin doğruluğunu tespit etmek amacıyla çapraz doğrulama yöntemleri geliştirilmiştir. Büyük veri kümeleri kullanılarak sınıflandırma performansı yüksek modeller oluşturulabilmektedir. Çapraz doğrulama yöntemlerini uygulamanın amacı ise küçük veri kümelerinde eğitilen modelin aşırı öğrenmesini engellemektir.

Birini Dışarıda Bırakma çapraz doğrulama yöntemi az sayıda veriye sahip veri kümelerinde sık kullanılan yöntemlerden biridir. Veri kümesi içerisindeki her veri, test amacıyla bir defa seçilirken, diğer tüm verilerin eğitim kümesine dahil olmasıyla model eğitilir ve test edilir. Her veri için bu işlemin tekrarlanmasıyla doğruluk değerleri elde

edilir. Doğruluk değerlerinin ortalamasının alınmasıyla modelin doğruluk performansı hesaplanır.



Şekil 3.8. Birini dışarda bırakma çapraz doğrulama yöntemi.

3.8. Değerlendirme Ölçütleri

Makine öğreniminde oluşturulan modelin sınıflandırma performansını ölçmek amacıyla değerlendirme ölçütleri kullanılmaktadır. Sınıflandırması yapılan verilerden kaç tanesinin doğru ya da yanlış sınıfa ait olduğu karmaşıklık matrisi ile incelenerek modelin başarısı hesaplanır. Karmaşıklık matrisinde verinin gerçekten var olduğu sınıfın doğru ya da yanlış tahminlerinin yer aldığı değerler bulunmaktadır. Bu çalışmada pozitif değerler hasta, negatif değerler sağlıklı sınıfı oluşturmaktadır. Hasta katılımcılar hasta olarak tahmin edilmişse gerçek pozitif (GP), yanlış tahmin edilmişse yanlış negatif (YN); sağlıklı katılımcılar sağlıklı olarak tahmin edilmişse gerçek negatif (GN), hasta olarak tahmin edilmişse yanlış pozitif (YP) olarak adlandırılır. Karmaşıklık matrisindeki değerler toplamı veri kümesinde yer alan katılımcı sayılarının toplamını vermektedir. Doğruluk, hassaslık, kesinlik, f-ölçütü ve Matthews korelasyon katsayısı gibi parametreler hesaplanarak modelin başarısı hakkında yorum yapılmaktadır.

Çizelge 3.3. Karmaşıklık matrisi.

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

3.8.1. Doğruluk

Elde edilen sınıflandırma sayıları içerisinde doğru tahmin edilen katılımcı sayısının tüm veri kümesindeki katılımcı sayısında oranıdır. Bu çalışmada, doğru sınıflarda tahmin edilen Parkinson hastalarının ve sağlıklı katılımcıların değerini göstermektedir.

$$Doğruluk = \frac{GP + GN}{GP + YP + GN + YN} \tag{3.11}$$

3.8.2. Duyarlılık

Modelin pozitif sınıfta gerçekleşen sınıflandırmaları ne kadar uygun şekilde gerçekleştirdiğini ölçmektedir. Bu çalışmada, hasta katılımcıların tüm hasta sayısına oranı şeklinde ifade edilmektedir.

$$Duyarlılık = \frac{GP}{GP + YN} \tag{3.12}$$

3.8.3. Kesinlik

Modelin pozitif olarak yorumladığı durumları ne kadar iyi ölçtüğünü göstermektedir. Hasta katılımcıların, kaç tanesinin gerçekten hasta olduğunun oranını vermektedir.

$$Kesinlik = \frac{GP}{GP + YP} \tag{3.13}$$

3.8.4. F-ölçütü

Hassaslık ve kesinliğin ölçümünün yeterli olmadığı durumlarda harmonik ortalama olarak da adlandırılan f-ölçütü, verilerin doğruluğunun ölçümüdür. F-ölçütü, hassaslık ve kesinliğin harmonik ortalaması alınarak hesaplanır.

$$f_ölçütü = 2\,x\,\frac{Duyarlılık\ x\ Kesinlik}{Duyarlılık + Kesinlik} \tag{3.14}$$

3.8.5. Matthews Korelasyon Katsayısı

MKK, dengesiz veri kümesinden etkilenmeyen ve sınıflandırmanın rastgeleliği hakkında bilgi veren bir ölçüttür. [-1, +1] aralığında değerler alarak sınıflar arası korelasyonu belirler. MKK değeri +1'e yaklaştıkça modelin doğrulunda artış gözlemlenmektedir. -1'e yaklaştıkça yanlış bir sınıflandırmadan söz edilebilirken, 0 olduğu değerler modelin rastgele bir sınıflandırma olduğunu göstermektedir.

$$MKK = \frac{(GPxGN)(YNxYP)}{\sqrt{(GP+YN)(GP+GP)(GN+YN)(GN+YP)}} \quad (3.15)$$

BÖLÜM 4

DENEYSEL SONUÇLAR

Bu tez çalışmasında, Parkinson hastalarına ve sağlıklı katılımcılara ait tuş vuruş verileri incelenerek hasta-sağlıklı ayrımı yapabilen bir modelin geliştirilmesi amaçlanmıştır. Alt veri kümeleri oluşturularak normalizasyon yöntemlerinin başarısı ve öznitelik seçim algoritmalarının performansları, (i) ilaç kullanım durumlarına, (ii) hastalık şiddetine ve (iii) cinsiyete göre incelenmiştir.

Donma, gecikme ve uçuş sürelerine ait tuş vuruş yönlerinden her biri için 14 öznitelik oluşturularak, toplamda 378 öznitelik (*Ek 1*) elde edilmiştir. Çıkarılan öznitelikler k-En Yakın Komşu (k-EYK), Destek Vektör Makineleri (DVM) ve Rastgele Orman (RO) sınıflandırma algoritmalarına verilerek 14 farklı veri kümesinde ilaç kullanım durumlarına, hastalık şiddetine ve cinsiyete göre başarı performansları incelenmiştir. Ayrıca Birini Dışarda Bırak Çapraz Doğrulama Yöntemi uygulanarak aşırı öğrenmenin önüne geçilmeye çalışılmıştır.

k-EYK, DVM sınıflandırıcılarında bütün parametre değerleri tek tek incelenmiş ve en iyi sonucu veren parametreler seçilmiştir. k-EYK algoritmasında k=3,5,7,9 komşuluk değerleri ve Euclid, City-Block, Chebychev ve Minkowski uzaklık ölçümleri; DVM algoritmasında ise lineer (t=0), polinom (t=1), RBF (t=2), sigmoid (t=3) çekirdek fonksiyonları ayrı ayrı değerlendirilmiştir. Ayrıca RO algoritmasının yapısından kaynaklanan rastgelelikten dolayı model 100 defa çalıştırılarak elde edilen doğruluk değerlerinin ortalaması alınmıştır.

Öznitelik seçim yöntemleri (mRmR, Relief, SFS, RO) uygulanarak en iyi performansı gösteren ilk 100 öznitelik belirlenmiş ve bu özniteliklere ait sınıflandırma sonuçları çalışmaya dahil edilmiştir. Ayrıca tuş vuruşunun el-parmak hareketlerinde meydana gelen bozulmaları inceleyebilmek amacıyla, 4 farklı öznitelik seçim yönteminde de seçilen ilk 20 öznitelik içerisinden en az 3 yöntemde ortak olan 5 öznitelik belirlenmiş ve sınıflandırıcı sonuçları elde edilerek öznitelik grupları ve tuş vuruş yönleri incelenmiştir.

4.1. Veri Kümesi

Bu tez çalışması kapsamında, Physionet veri tabanından alınan veri kümesi içerisinde, her katılımcıya ait bir günden fazla veri bulunurken, veri içerisindeki tuş vuruşları bazı katılımcılarda oldukça az sayıdaydı. Bundan dolayı, her katılımcının tek günlük verileri içerisinden en az 2000 tuş vuruşu içeren veriler çalışmaya dahil edilmiştir. Veri seti içerisinde katılımcılar tarafından doldurulan sağlık bilgileri dikkate alınarak, 14 farklı veri kümesi oluşturulmuştur. Çizelge 4.1'deki veri kümelerinin her birine belirtilen tüm işlem basamakları uygulanarak sınıflandırma doğruluk performansları ve seçilen özniteliklerin önemi incelenmiştir.

Veri kümesi 1, ilaç kullanım durumu, hastalık şiddeti veya cinsiyet bilgilerine bakılmaksızın, en az 2000 tuş vuruşu yapan tüm katılımcılardan meydana gelen en büyük veri kümesidir. Veri kümesi 2-5, veri kümesi 1'deki dengesiz katılımcı sayısının önüne geçmek ve ilaç kullanım durumlarını incelemek amacıyla oluşturulmuştur. Veri kümesi 6-12'de hastalık şiddetinin sınıflandırma başarısı üzerindeki etkisini incelemek için oluşturulmuştur. Veri kümesi 13 ve 14 ise modelin cinsiyet üzerindeki performansı araştırmak için oluşturulmuştur.

Çizelge 4.1. Oluşturulan veri kümeleri.

	Sağlıklı	Parkinson
Veri Kümesi 1	Tüm Sağlıklı (30 kişi)	Tüm Parkinson hastaları (88 kişi)
Veri Kümesi 2	Tüm Sağlıklı (30 kişi)	Tüm Parkinson hastaları içerisinden seçilen rastgele 30 kişi
Veri Kümesi 3	Tüm Sağlıklı (30 kişi)	Herhangi bir ilaç kullanan Parkinson hastaları (78 kişi)
Veri Kümesi 4	Tüm Sağlıklı (30 kişi)	Herhangi bir ilaç kullanan Parkinson hastaları içerisinden seçilen rastgele 30 kişi
Veri Kümesi 5	Tüm Sağlıklı (30 kişi)	Herhangi bir ilaç kullanmayan Parkinson hastaları (10 kişi)
Veri Kümesi 6	Tüm Sağlıklı (30 kişi)	Hafif derecede Parkinson hastaları (37 kişi)
Veri Kümesi 7	Tüm Sağlıklı (30 kişi)	Hafif derecede Parkinson hastaları içerisinden seçilen rastgele 10 kişi
Veri Kümesi 8	Tüm Sağlıklı (30 kişi)	Hafif derece ve herhangi bir ilaç kullanan Parkinson hastaları (30 kişi)
Veri Kümesi 9	Tüm Sağlıklı (30 kişi)	Hafif derece ve herhangi bir ilaç kullanmayan Parkinson hastaları (7 kişi)
Veri Kümesi 10	Tüm Sağlıklı (30 kişi)	Orta derece Parkinson hastaları (39 kişi)
Veri Kümesi 11	Tüm Sağlıklı (30 kişi)	Orta derecede Parkinson hastaları içerisinden seçilen rastgele 10 kişi
Veri Kümesi 12	Tüm Sağlıklı (30 kişi)	Ağır derece Parkinson hastaları (10 kişi)
Veri Kümesi 13	Sağlıklı kadın (16 kişi)	Parkinson hastası kadın (51 kişi)
Veri Kümesi 14	Sağlıklı erkek (14 kişi)	Parkinson hastası erkek (37 kişi)

4.2. Normalizasyon Yöntemlerinin Etkinliği

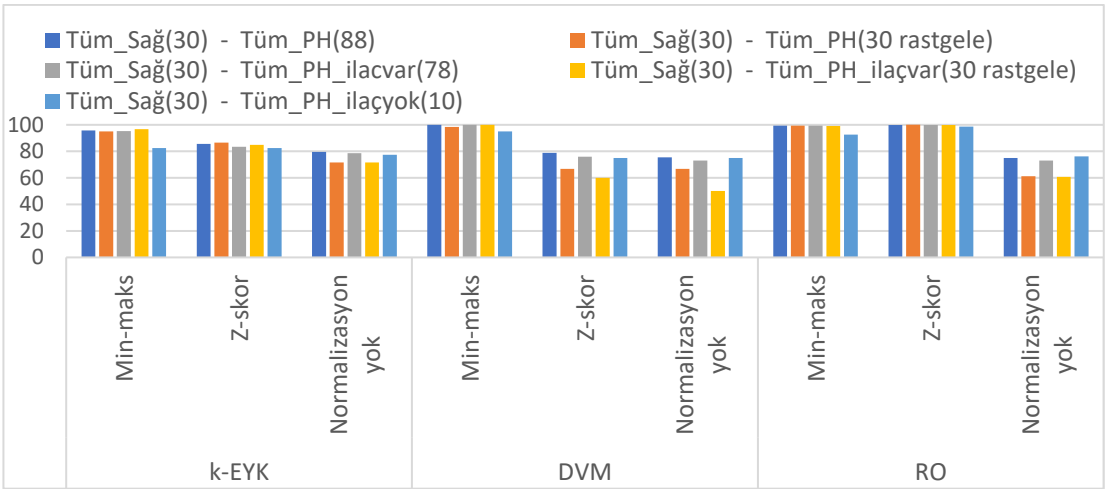
Çalışmanın ilk aşamasında, Parkinson hastaları ve sağlıklı katılımcıların günlük tuş vuruş verileri içerisinden en az 2000 tuş vuruşu içeren veriler dahil edilmiştir. Veriyi gürültüden arındırmak amacıyla, veri içerisinde yer alan hatalı satırlar tek tek düzenlenmiş, sadece katılımcı ismi yer alan veri satırları ise silinerek normalizasyon işleminde kullanılmak üzere hazır hale getirilmiştir. Veride yer alabilecek uç değerlerin

sınıflandırma performansını etkilememesi için, minimum-maksimum (-1 +1) ve z-skor normalizasyon yöntemleri kullanılarak veri belirli bir aralığa sıkıştırılmıştır. Normalizasyon işleminin yapılmadığı ve 2 farklı normalizasyon işleminin uygulandığı durumlarda elde edilen veriler ile öznitelik çıkarma işlemleri uygulanmıştır. Çizelge 4.2.'de oluşturulan veri kümelerinin uygulanan normalizasyon yöntemlerine göre, doğruluk değerleri yer almaktadır.

Çizelge 4.2. Normalizasyon yöntemlerine göre sınıflandırma başarı performansları.

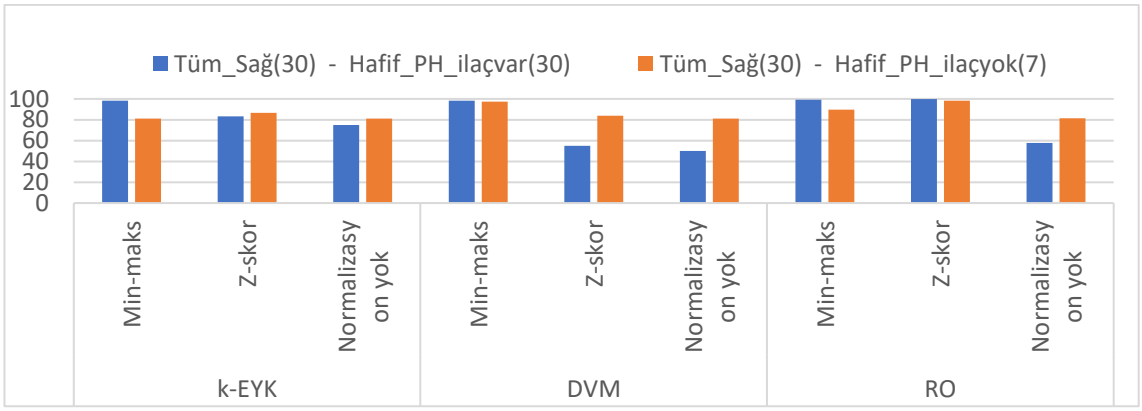
378 Öznitelik (Doğruluk)		k-EYK			DVM			RO		
		Min-maks	Z-skor	Normalizasyon yok	Min-maks	Z-skor	Normalizasyon yok	Min-maks	Z-skor	Normalizasyon yok
1	**Tüm_Sağlıklı (30) - Tüm_PH (88)**	**95,76**	85,59	79,66	**100**	78,81	75,42	99,34	**99,93**	74,97
2	**Tüm_Sağlıklı (30) - Tüm_PH (30 rastgele)**	**95**	86,66	71,66	**98,33**	66,66	66,66	99,30	**99,98**	61,30
3	**Tüm_Sağlıklı(30) - Tüm_PH_ilacvar (78)**	**95,37**	83,33	78,70	**100**	75,92	73,14	99,46	**99,95**	72,94
4	**Tüm_Sağlıklı(30) - Tüm_PH_ilaçvar (30 rastgele)**	**96,66**	85	71,66	**100**	60	50	99,07	**99,96**	60,75
5	**Tüm_Sağlıklı(30) - Tüm_PH_ilaçyok (10)**	**82,50**	**82,50**	77,50	**95**	75	75	92,55	**98,60**	76,15
6	**Tüm_Sağlıklı (30) - Hafif_PH (37)**	**98,50**	79,10	76,11	**100**	58,20	56,71	99,34	**99,91**	57,81
7	**Tüm_Sağlıklı (30) - Hafif_PH (10 rastgele)**	**87,50**	**82,50**	80	**92,50**	75	75	93,80	**99,92**	75,92
8	**Tüm_Sağlıklı (30) - Hafif_PH_ilaçvar (30)**	**98,33**	83,33	75	**98,33**	55	50	99,08	**99,92**	57,60
9	**Tüm_Sağlıklı (30) - Hafif_PH_ilaçyok (7)**	81,08	**86,48**	81,08	**97,29**	83,78	81,08	89,62	**98,24**	81,43
10	**Tüm_Sağlıklı (30) - Orta_PH (39)**	**94,20**	79,71	73,91	**100**	57,97	57,97	99,39	**99,75**	58,55
11	**Tüm_Sağlıklı (30) - Orta_PH (10 rastgele)**	82,50	**90**	70	**90**	75	75	93,70	**99,95**	77,32
12	**Tüm_Sağlıklı (30) - Ciddi_PH (10)**	**87,50**	80	77,50	**90**	75	75	92,77	**99,52**	75,58
13	**Kadın_Sağlıklı (16) - Kadın_PH (51)**	**88,59**	80,59	82,08	**98,50**	76,11	77,61	98,58	**99,96**	76,43
14	**Erkek_Sağlıklı (14) - Erkek_PH (37)**	**88,23**	80,39	76,47	96,07	**98,03**	72,54	96,25	**98,92**	73,63

Tüm Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümeleri (veri kümesi 1-5) incelendiğinde, veri kümesi 1'de Parkinson ve sağlıklı katılımcı sayılarındaki dengesizliğin önüne geçmek amacıyla veri kümesi 2'de Parkinson hastaları arasından rastgele seçilen 30 kişi oluşturulmuş ve elde edilen sonuçlarda anlamlı bir farklılık gözlemlenmemiştir. Ancak minimum-maksimum normalizasyon yöntemi ile DVM sonuçlarında elde edilen %100 doğruluk değeri, katılımcı sayılarının dengelenmesiyle %98,33 doğruluk değerini elde etmiştir. k-EYK algoritmasında minimum-maksimum normalizasyon yöntemi ile elde edilen doğruluk değerinde, %95,76'dan %95'e düşüş meydana gelmiştir. RO'da ise önceden z-skor normalizasyon yöntemi ile %99,93 olan doğruluk değeri katılımcı sayılarının dengelenmesiyle %99,98 olmuştur. Veri kümesi 3 ve 4'te ise ilaç kullanan Parkinson hastalarında katılımcı sayıları dengelenerek, başarı performansında artışlar beklenmiştir. Ancak hem minimum-maksimum normalizasyon hem de z-skor yöntemleri ile elde edilen sonuçlarda anlamlı farklılıklar gözlemlenmemiştir (İlaç kullanan tüm Parkinson'da doğruluk değerleri %95,37-100 iken, rastgele seçilen ilaç kullanan Parkinson'da %96,66-100). k-EYK ve DVM'de minimum-maksimum normalizasyon yöntemi, RO'da ise z-skor yöntemi en başarı normalizasyon yöntemi olmuştur. Tüm Parkinson hastalarında ilaç kullanım durumları incelendiğinde, minimum-maksimum normalizasyon yöntemi ile elde edilen sonuçlarda, model ilaç kullanan Parkinson hastalarını %95,37-100 doğruluk ile ayırırken ilaç kullanmayan Parkinson hastalarında %82,50-95 başarı performansı göstermiştir. Tüm Parkinson hastaları ve sağlıklı katılımcılar ile oluşturulan veri kümeleri incelendiğinde, k-EYK ve DVM algoritmalarında minimum-maksimum normalizasyon yöntemi hasta-sağlıklı ayrımında en başarılı yöntem olduğu görülmüştür. RO algoritmasında ise z-skor en yüksek sonuçları verirken, minimum-maksimum normalizasyon yöntemiyle aralarında anlamlı bir farklılık gözlemlenmemiştir (Şekil 4.1).



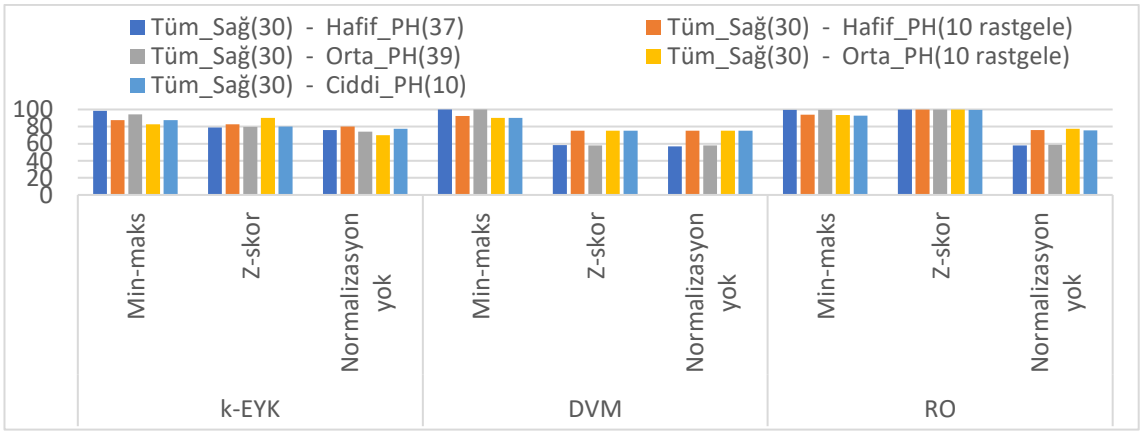
Şekil 4.1. Tüm Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları.

Tüm Parkinson hastaları ile yapılan sınıflandırmalarda modelin ilaç kullanmayan hastaları, ilaç kullananlara göre daha yüksek doğrulukla ayırması beklenmiştir. Ancak tam tersi sonuçlar elde edilmiştir. Tüm Parkinson hastaları içerisinde hastalık şiddetlerinin bilinmemesinden dolayı hafif derece Parkinson hastaları da ilaç kullanım durumlarına göre incelenmiştir. Orta ve ciddi derece Parkinson hastalarını içeren veri kümesindeki katılımcıların neredeyse hepsi ilaç kullandığından dolayı bu veri kümelerinin ilaç kullanım durumları incelenmemiştir. Çizelge 4.2.'ye göre, ilaç kullanan (veri kümesi 8) ve kullanmayan (veri kümesi 9) Parkinson hastaları ile sağlıklı katılımcılar arasında yapılan sınıflandırma sonuçlarında ilaç kullanan hastaların daha yüksek doğrulukla sağlıklı katılımcılardan ayrıldığı görülmüştür. k-EYK'de, ilaç kullananlarda minimum-maksimum normalizasyon %98,33, kullanmayanlarda z-skor normalizasyon %86,48; DVM'de, minimum-maksimum normalizasyon ilaç kullananlarda %98,33, kullanmayanlarda %97,29; RO'da ise z-skor normalizasyon ilaç kullananlarda %99,92, kullanmayanlarda ise %98,24 başarı performansları elde edilmiştir. Şekil 4.2'de verilen grafik incelendiğinde, z-skor normalizasyon RO algoritmasında, minimum-maksimum normalizasyon ise k-EYK ve DVM algoritmalarında daha başarılıdır. Hafif derece Parkinson hastaları belirti göstermedikleri takdirde ilaç kullanmamaktadırlar. İlaç kullanan hastalarda ise ilaç her ne kadar belirtileri ortadan kaldırsa da model bu veri kümesinde hastaları daha iyi tespit edebilmiştir.



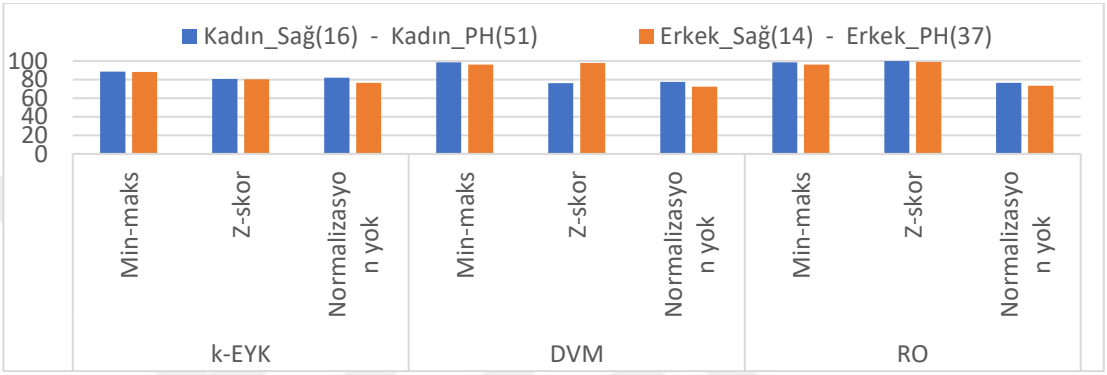
Şekil 4.2. Hafif derece Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları.

Hastalık şiddetinin başarı performansını nasıl etkilediğini araştırmak amacıyla, hafif, orta ve ciddi Parkinson hastalarından oluşan veri kümeleri incelenmiştir (veri kümesi 6, 7, 10-12). Ciddi derece Parkinson hasta veri kümesinde katılımcı sayısının az olmasından dolayı hafif ve orta derece Parkinson hastaları içerinden rastgele seçilen 10 kişi seçilerek ciddi Parkinson hastalarıyla kıyaslanmıştır. Çizelge 4.2' de elde edilen sonuçlara göre, hastalık şiddetinin artmasıyla modelin sağlıklı katılımcılardan daha yüksek doğrulukla ayrılması beklenmiştir. Sağlıklı katılımcılar ile hafif derece Parkinson hastalığı %82,5-99,92; orta derece Parkinson hastaları %90-99,95; ciddi derece Parkinson hastalığı ise %87,5-99,52 doğruluk değerleri ile ayrılmıştır. Şekil 4.3'te de görüldüğü gibi minimum-maksimum normalizasyon hem k-EYK hem de DVM'de en başarılı normalizasyon yöntemi iken, z-skor normalizasyon ile RO'da daha yüksek değerler elde edilmiştir (Şekil 4.3).



Şekil 4.3. Hastalık şiddetine göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları.

Parkinson hastaları ve sağlıklı katılımcılar cinsiyetlerine göre ayrılarak farklı sınıflandırıcılardaki normalizasyon yöntemlerinin başarı performansları incelenmiştir. Kadın katılımcılar ile yapılan sınıflandırma sonuçlarında (%88,59-99,96), erkek katılımcılara (%88,23-98,92) göre daha yüksek doğruluk değerleri elde edilmiştir. Kadın katılımcılarda, minimum-maksimum normalizasyon, erkek katılımcılarda ise z-skor normalizasyon en başarılı normalizasyon yöntemleri olmuştur. (Şekil 4.4).



Şekil 4.4. Cinsiyete göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde normalizasyon yöntemlerinin başarı performansları.

Oluşturulan veri kümeleri incelendiğinde k-EYK ve DVM algoritmalarında minimum-maksimum normalizasyon yönteminin hasta-sağlıklı ayrımında en başarılı normalizasyon yöntemi olduğu görülmektedir. RO algoritmasında ise z-skor en başarılı normalizasyon yöntemi iken minimum-maksimum yöntemi ile elde edilen sonuçlarla karşılaştırıldığında anlamlı bir farklılığın olmadığı gözlemlenmiştir. Bu aşamadan sonra yapılan işlemlerde minimum-maksimum normalizasyon yöntemi kullanılmıştır.

4.3. Öznitelik Seçim Yöntemlerinin Etkinliği

Sınıflandırma performansını etkileyen etmenlerden biri de öznitelik sayısıdır. Fazla sayıda öznitelik içeren modellerde, doğruluk değerlerini düşürebilecek öznitelikler de yer alabilmektedir. 378 öznitelik yerine en başarılı olan özniteliklerin kullanılmasıyla modelin daha başarılı bir performans göstermesi amaçlanmıştır.

Klavye tuş vuruş verilerinden elde edilen özniteliklere yaygın olarak kullanılan mRmR, Relief, SFS ve RO öznitelik seçim yöntemleri uygulanarak en başarılı ilk 100 öznitelik her yöntemde belirlenmiştir. Belirlenen 100 öznitelik ilk 10, ilk 20, …, ilk 100

olacak şekilde gruplandırılmış ve 3 sınıflandırma algoritmasındaki sonuçları 378 öznitelikte olduğu gibi incelenmiş ve veri kümelerinde en iyi sonuçları veren öznitelik seçim algoritmaları belirlenmiştir. 14 farklı veri kümesinde belirlenen ilk 20 öznitelik Çizelge 4.3'te elde edilen sınıflandırma sonuçları da Çizelge 4.4'te verilmiştir.

Çizelge 4.3. Öznitelik seçim yöntemleri ile seçilen ilk 20 öznitelik.

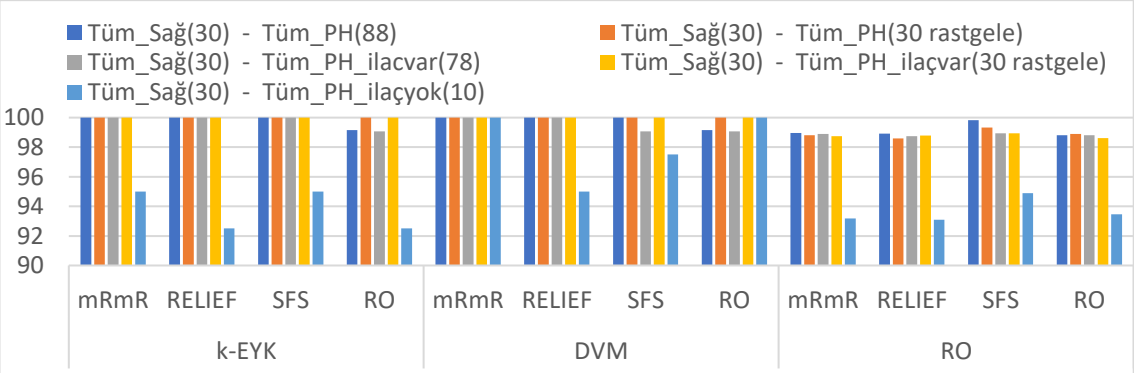
Seçilen İlk 20 Öznitelik		Yöntem	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Tüm_Sağlıklı (30) - Tüm_PH (88)	mRmR	208	205	88	48	377	87	164	44	335	321	36	19	276	90	251	47	86	332	41	153
		Relief	208	87	88	44	48	90	86	377	36	47	19	89	164	166	276	73	290	335	360	262
		SFS	88	14	178	22	23	37	50	41	62	71	20	145	51	58	153	31	16	65	120	78
		RO	208	34	88	89	87	33	75	44	48	76	90	36	19	64	377	37	72	30	16	65
2	Tüm_Sağlıklı (30) - Tüm_PH (30 rastgele)	mRmR	13	87	12	153	36	48	167	360	321	237	44	90	41	68	332	86	265	279	340	5
		Relief	87	48	44	153	36	47	90	86	167	41	68	12	360	89	164	169	276	37	262	290
		SFS	87	72	71	6	76	23	78	79	69	9	103	139	29	256	36	1	2	11	3	16
		RO	48	37	86	153	87	90	5	89	44	64	36	47	65	68	23	34	51	50	167	41
3	Tüm_Sağlıklı (30) - Tüm_PH_ilaçvar (78)	mRmR	208	191	87	48	88	377	164	44	335	36	321	19	332	90	41	47	85	86	251	276
		Relief	208	87	88	44	48	90	86	377	36	47	19	164	89	166	290	276	73	335	360	262
		SFS	88	14	262	27	4	20	16	38	22	50	72	23	101	2	31	58	107	118	76	120
		RO	87	33	44	47	90	30	208	88	75	34	377	51	89	86	251	20	72	36	6	48
4	Tüm_Sağlıklı (30) - Tüm_PH_ilaçvar (30 rastgele)	mRmR	68	44	87	164	41	167	321	36	153	48	237	12	340	90	360	86	332	47	27	169
		Relief	87	48	44	41	47	90	167	36	86	153	12	89	169	164	360	37	262	321	276	340
		SFS	90	36	2	45	183	65	72	118	89	76	78	9	225	3	86	101	6	4	30	79
		RO	86	37	89	87	90	47	36	44	100	64	12	23	104	48	51	332	214	50	34	2
5	Tüm_Sağlıklı (30) - Tüm_PH_ilaçyok (10)	mRmR	100	44	321	78	5	121	169	104	79	65	251	8	103	127	195	2	153	76	125	6
		Relief	100	5	104	78	79	2	6	103	169	121	76	251	65	72	127	15	8	321	195	41
		SFS	78	13	79	16	72	20	23	50	22	36	321	89	47	167	125	292	54	120	30	51
		RO	78	79	5	121	104	2	100	321	153	6	76	72	65	82	103	346	41	9	73	334
6	Tüm_Sağlıklı (30) - Hafif_PH (37)	mRmR	12	44	118	78	321	279	335	48	5	293	72	106	340	167	114	19	111	76	47	41
		Relief	48	44	78	72	76	47	5	19	321	111	293	335	167	114	118	103	41	106	340	279
		SFS	9	16	13	1	22	21	8	29	17	126	307	47	12	15	20	23	24	50	32	51
		RO	78	20	5	16	6	48	47	72	76	79	103	2	118	23	114	44	50	19	51	340
7	Tüm_Sağlıklı (30) - Hafif_PH (10 rastgele)	mRmR	88	12	72	23	363	121	5	79	1	20	153	76	68	44	37	78	16	127	100	237
		Relief	79	23	72	78	76	121	5	12	100	104	68	237	120	99	169	127	1	16	363	20
		SFS	79	72	12	1	14	16	2	3	4	20	17	30	8	103	6	19	27	5	13	9
		RO	121	79	23	100	16	1	72	104	12	47	51	321	44	5	6	114	127	76	20	118
8	Tüm_Sağlıklı (30) - Hafif_PH_ilaçvar (30)	mRmR	12	44	72	114	106	293	5	335	48	78	118	167	340	279	19	76	284	103	47	111
		Relief	72	76	48	44	5	78	47	111	293	114	118	106	19	167	335	103	41	340	363	50
		SFS	72	78	58	15	79	269	113	2	9	23	22	33	50	86	68	34	8	100	99	278
		RO	76	79	48	78	44	72	6	118	103	23	20	47	16	5	293	51	335	19	2	14
9	Tüm_Sağlıklı (30) - Hafif_PH_ilaçyok (7)	mRmR	20	365	100	79	65	16	321	76	5	121	104	72	23	253	8	78	195	153	2	103
		Relief	72	76	20	16	5	79	65	100	104	78	253	309	2	103	6	121	120	22	114	251
		SFS	41	195	78	186	64	22	267	95	172	13	221	271	237	113	130	185	321	69	72	281
		RO	2	72	121	100	16	76	9	20	8	321	114	279	104	79	145	6	162	253	103	106
10	Tüm_Sağlıklı (30) - Orta_PH (39)	mRmR	208	88	69	377	48	36	87	321	44	103	90	251	175	335	86	41	293	22	138	12
		Relief	208	88	87	377	48	44	36	90	86	69	138	166	41	321	89	251	37	341	64	195
		SFS	88	2	37	59	36	1	3	321	41	4	40	9	18	6	7	13	29	27	23	32
		RO	208	88	90	86	37	377	44	5	36	89	117	50	64	87	48	65	103	23	6	2
11	Tüm_Sağlıklı (30) - Orta_PH (10 rastgele)	mRmR	321	87	167	103	195	36	279	44	163	55	5	208	75	169	104	56	90	377	57	22
		Relief	87	167	103	50	100	104	321	41	163	253	169	107	5	195	55	377	208	57	36	152
		SFS	65	79	121	64	23	58	16	62	201	20	72	76	78	51	87	75	279	22	41	27
		RO	86	87	167	6	89	90	100	64	37	321	195	36	1	103	40	279	2	5	293	302
12	Tüm_Sağlıklı (30) - Ciddi_PH (10)	mRmR	23	76	153	8	44	22	1	78	20	164	120	278	79	139	16	65	72	99	30	87
		Relief	22	23	78	72	79	76	153	120	20	16	65	99	121	57	43	64	237	106	1	8
		SFS	23	1	164	11	40	55	15	79	7	239	29	43	57	21	27	39	99	71	82	113
		RO	5	23	22	20	79	76	262	377	48	2	248	6	153	100	44	78	121	104	319	57
13	Kadın_Sağlıklı (16) - Kadın_PH (51)	mRmR	27	36	117	136	88	205	87	19	276	166	321	5	44	89	262	100	169	90	290	335
		Relief	87	88	36	89	90	19	86	166	44	48	136	234	117	192	27	276	164	169	104	341
		SFS	87	9	20	36	3	106	27	11	8	10	16	22	50	19	21	35	23	31	28	58
		RO	2	88	37	5	89	87	86	36	6	90	51	111	19	360	44	50	30	23	104	321
14	Erkek_Sağlıklı (14) - Erkek_PH (37)	mRmR	208	362	363	87	44	251	16	247	155	48	377	153	86	293	50	41	20	167	211	75
		Relief	208	87	44	48	377	251	47	86	75	90	247	89	293	211	153	16	20	50	167	62
		SFS	87	33	72	9	30	18	34	41	373	74	3	106	78	104	76	79	369	37	83	200
		RO	65	47	30	51	33	72	208	75	76	44	73	34	86	37	87	50	377	62	293	58

Çizelge 4.4. Öznitelik seçim yöntemlerine göre sınıflandırma başarı performansları.

Seçilen Öznitelikler (Doğruluk)		k-EYK				DVM				RO			
		mRmR	Relief	SFS	RO	mRmR	Relief	SFS	RO	mRmR	Relief	SFS	RO
1	**Tüm_Sağlıklı (30) - Tüm_PH (88)**	**100**	**100**	**100**	99,15	**100**	**100**	**100**	99,15	98,97	98,92	**99,82**	98,80
2	**Tüm_Sağlıklı (30) - Tüm_PH (30 rastgele)**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	98,82	98,60	**99,33**	98,90
3	**Tüm_Sağlıklı(30) - Tüm_PH_ilacvar (78)**	**100**	**100**	**100**	99,07	**100**	**100**	99,07	99,07	98,90	98,75	**98,93**	98,81
4	**Tüm_Sağlıklı(30) - Tüm_PH_ilaçvar (30 rastgele)**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	98,75	98,78	**98,95**	98,62
5	**Tüm_Sağlıklı(30) - Tüm_PH_ilaçyok (10)**	**95**	92,50	**95**	92,50	**100**	95	97,50	**100**	93,18	93,10	**94,88**	93,47
6	**Tüm_Sağlıklı (30) - Hafif_PH (37)**	**100**	**100**	**100**	**100**	**100**	**100**	97,01	98,50	**98,84**	**98,84**	98,43	98,79
7	**Tüm_Sağlıklı (30) - Hafif_PH (10 rastgele)**	**100**	97,50	97,50	97,50	**97,50**	**97,50**	95	**97,50**	94,20	94,55	**94,60**	94,53
8	**Tüm_Sağlıklı (30) - Hafif_PH_ilaçvar (30)**	**100**	**100**	**100**	**100**	**100**	**100**	98,33	98,33	98,05	98,07	**98,85**	98,18
9	**Tüm_Sağlıklı (30) - Hafif_PH_ilaçyok (7)**	97,30	91,89	94,59	**97,33**	**100**	97,29	94,59	97,29	92,86	92,65	93,38	**93,57**
10	**Tüm_Sağlıklı (30) - Orta_PH (39)**	**100**	98,55	**100**	98,55	**98,55**	**98,55**	**98,55**	**98,55**	98,81	98,84	**99,33**	98,84
11	**Tüm_Sağlıklı (30) - Orta_PH (10 rastgele)**	95	92,50	95,20	**97,50**	**100**	**100**	95	97,50	96,07	95,50	96,03	**96,62**
12	**Tüm_Sağlıklı (30) - Ciddi_PH (10)**	**100**	97,50	97,50	97,50	**97,50**	**97,50**	**97,50**	95	94,60	94,80	**96,40**	94,95
13	**Kadın_Sağlıklı (16) - Kadın_PH (51)**	**100**	98,51	**100**	98,51	**100**	98,50	98,50	98,50	97,81	97,91	**98,57**	97,97
14	**Erkek_Sağlıklı (14) - Erkek_PH (37)**	**100**	98,04	98,04	98,04	**100**	98,03	98,03	96,07	95,59	95,51	**96,10**	95,61

Tüm Parkinson hastaları ve sağlıklı katılımcılar ile oluşturulan veri kümelerinde (veri kümesi 1-5), tüm Parkinson hastaları ve ilaç kullanan hastalarda veri kümelerini dengeli hale getirmek sınıflandırma sonuçlarını anlamlı ölçüde etkilememiştir (Çizelge 4.4). Veri kümesi 1 ve 2'de k-EYK ve DVM algoritmaları ile yapılan sınıflandırmalarda mRmR, Relief ve SFS yöntemleri %100 doğruluk değeri elde ederken, bu yöntemlerle seçilen ilk 20 öznitelik içerisinde boşluk tuşuna ait tuş vuruşlarının mRmR ve Relief'te

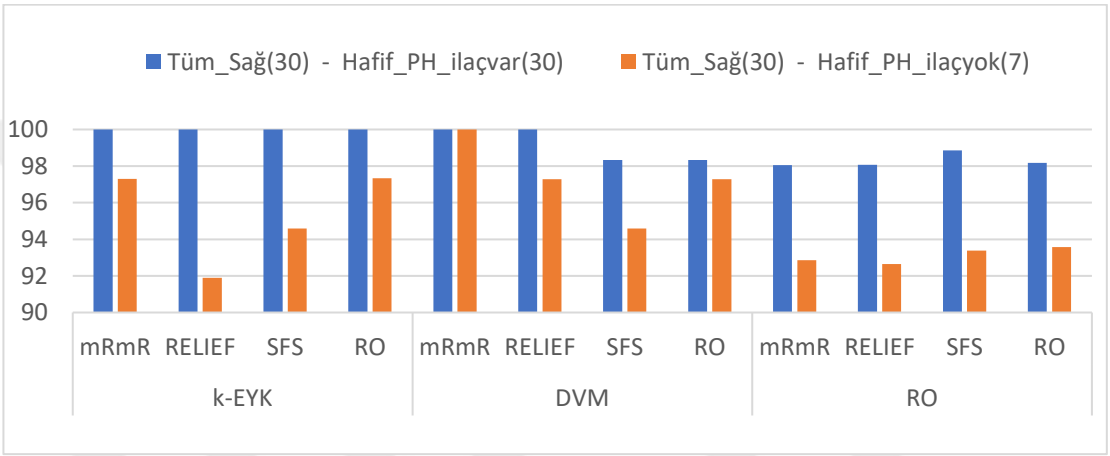
sağ ve sol yönde yapılan vuruşlara göre çoğunlukta olduğu görülmüştür. SFS'de ise sağ ve sol tuşların sayısı daha fazla olmuştur. RO sınıflandırma algoritmasında ise SFS öznitelik seçme algoritması ile seçilen öznitelikler %99,33-99,82 başarı performansını yakalamış ve bu özniteliklerin sağ ve sol tuşlara yapılan vuruşlardan meydana geldiği belirlenmiştir. Veri kümesi 3-5 incelendiğinde, model ilaç kullanan hastaları (%98,62-100 doğruluk) kullanmayanlara (%92,5-100 doğruluk) göre daha yüksek doğrulukla sağlıklı katılımcılardan ayırmayı başarmıştır. k-EYK ve DVM algoritmalarında, nerdeyse tüm yöntemler; RO algoritmasında ise SFS yöntemi en başarılı öznitelikleri seçerek en yüksek sonuçların elde edilmesini sağlamıştır. İlaç kullanan veri kümelerinde mRmR, Relief ve RO öznitelik seçim yöntemlerinde boşluk tuş vuruşu, SFS'de ise sağ tuş vuruşu en çok seçilen öznitelikler iken, ilaç kullanmayanlarda ise SFS öznitelik seçim yöntemi ile sağ tuş vuruşuna ait veriler, diğer yöntemlerde ise sol tuş vuruşuna ait veriler daha çok seçilmiştir. Tüm Parkinson hastaları ve sağlıklı ile elde edilen veri kümelerinin sınıflandırma sonuçlarına göre, tüm kümelerde mRmR yöntemi, DVM algoritması sonuçlarında en başarılı öznitelik seçim yöntemidir (Şekil 4.5).



Şekil 4.5. Tüm Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları.

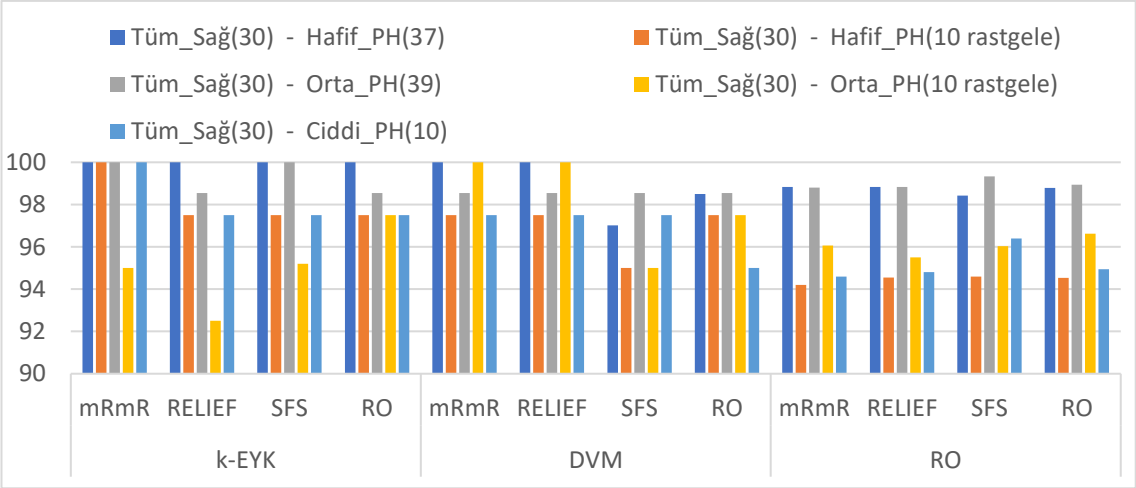
Tüm Parkinson hastalarıyla oluşturulan veri kümelerinde ilaç kullanmayan hastaların daha yüksek doğruluklarla sağlıklı katılımcılardan ayrılması beklenirken model ilaç kullanan hastaları daha yüksek doğrulukla ayırabilmiştir. Bunun, veri kümesindeki katılımcıların hastalık şiddetlerinin bilinmemesinden kaynaklandığı düşünülerek hafif derece Parkinson hastalarının ilaç kullanım durumları araştırılmıştır. Hafif derece Parkinson hastası veri kümelerine (veri kümesi 8, 9) ait 3 farklı sınıflandırıcıdaki sonuçlarda, model ilaç kullanan hastaları (%98,18-100 doğruluk)

kullanmayanlara (%92,65-100 doğruluk) göre daha iyi bir şekilde ayırabilmiştir. k-EYK'da, RO öznitelik seçim yöntemi; DVM'de, mRmR yöntemi en başarılı yöntem iken, RO algoritmasında öznitelik seçim yöntemleri ile elde edilen sonuçlarda büyük farklılıkların olmadığı görülmüştür. mRmR yöntemi ile seçilen ilk 20 öznitelik incelendiğinde, ilaç kullanan hastalarda boşluk tuşuna; kullanmayanlarda ise sol tuşuna yapılan vuruşların çoğunlukta olduğu belirlenmiştir. Şekil 4.6 incelendiğinde, mRmR algoritması DVM ile yapılan sınıflandırma sonuçlarında en yüksek doğruluk değerlerini elde etmiştir.



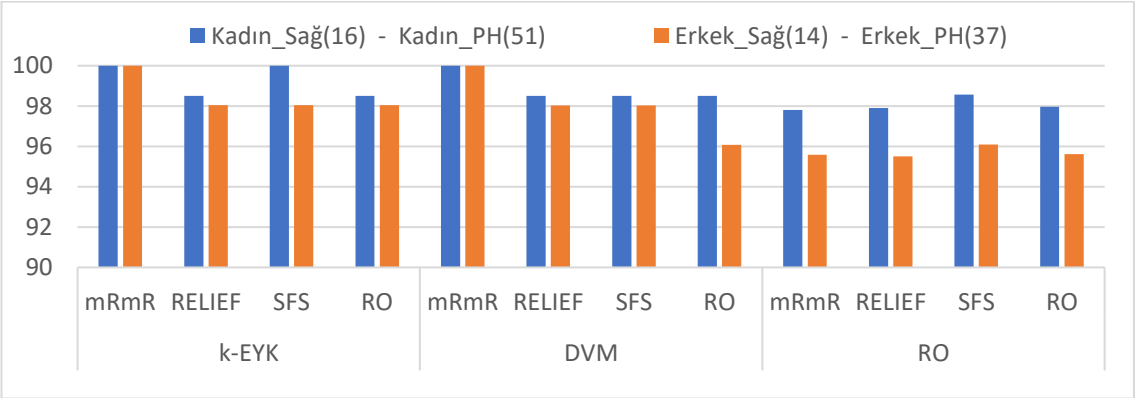
Şekil 4.6. Hafif derece Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları.

Veri kümeleri, ilaç kullanım durumlarına ek olarak hastalık şiddetine göre de ayrılmış (veri kümesi 6, 7, 10-12) ve hastalık şiddetinin artmasıyla doğruluk değerlerinin artması beklenmiştir. Ancak modelin başarı performans değerleri, rastgele seçilen veri kümelerinde hastalık şiddetinin artmasıyla artmamış hatta bazı veri kümelerinde düşüşler meydana gelmiştir. Çizelge 4.4'te yer alan değerler incelendiğinde, k-EYK ve DVM algoritmalarında en başarılı sonuçlar mRmR ve Relief yöntemleri ile elde edilirken, RO'da SFS yöntemi ile seçilen öznitelikler en yüksek sonuçları vermiştir. Şekil 4.7 hastalık şiddetine göre kullanılan öznitelik seçme yöntemleriyle elde edilen sınıflandırma sonuçlarının sütun grafiğidir. mRmR, Relief ve SFS yöntemleri ile seçilen öznitelikler içerisinde, hafif derece Parkinson'da sol; orta derecede boşluk tuşu; ciddi derecede ise sağ tuş vuruşuna ait özniteliklerin yer aldığı belirlenmiştir.



Şekil 4.7. Hastalık şiddetine göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları.

Cinsiyete göre ayrılarak elde edilen veri kümelerinde (veri kümesi 13, 14) ise k-EYK ve DVM algoritmaları mRmR ile seçilen; RO algoritması ise SFS yöntemi ile seçilen özniteliklerle daha yüksek doğruluk değerleri elde ederek hasta-sağlıklı ayrımı yapılabilmiştir (Şekil 4.8). mRmR öznitelik seçim yönteminde, boşluk tuş vuruşu; SFS yönteminde ise sağ tuş vuruşuna ait özniteliklerin sayısının oldukça fazla olduğu tespit edilmiştir.



Şekil 4.8. Cinsiyete göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde öznitelik seçim yöntemlerinin başarı performansları.

Farklı öznitelik seçim yöntemleri kullanılarak 14 farklı veri kümesinde ilaç kullanım durumuna, hastalık şiddetine ve cinsiyete göre incelen sınıflandırma sonuçlarında, mRmR yöntemi ile seçilen öznitelikler k-EYK ve DVM de en yüksek sonuçların elde edilmesini sağlarken, RO algoritmasında ise SFS yöntemi ile seçilen

öznitelikler hasta-sağlıklı ayrımını başarılı şekilde yapabilmiştir. Bu özniteliklerle seçilen öznitelikler incelendiğinde, veri kümelerinin özelliklerine göre tuş vuruş yönlerinin (sol, sağ ve boşluk) de değiştiği ancak boşluk tuş vuruşlarının genel olarak daha fazla olduğu gözlemlenmiştir. Literatürde yer alan bazı çalışmalarda boşluk tuşuna ait veri satırları çalışmaya dahil edilmezken, bu çalışmada boşluk tuş vuruşlarının dahil edilmesi ile boşluk tuşuna yapılan vuruşlar, Parkinson hastalarını sağlıklı katılımcılardan ayıran önemli bir özellik olmuştur.

4.4. Ortak Seçilen Özniteliklerle Performans Değerlendirmesi

Öznitelik çıkarma yöntemleri ile elde edilen tüm özniteliklerle (378 öznitelik) ve 4 farklı öznitelik seçim yöntemi (mRmR, Relief, SFS, RO) ile seçilen özniteliklerle (ilk 100 öznitelik) elde edilen sonuçlar incelenmiştir. Kullanılan normalizasyon yöntemlerinin ve öznitelik seçim yöntemlerinin veri kümelerindeki etkinliği araştırılmış ve sonuçlarda yer alan %100 doğruluk değerleri ile modelde aşırı öğrenme durumundan şüphelenilmiştir. Doğruluk değerinin %100 olması değerlendirme ölçütlerinin (f-ölçütü, MKK) de karşılaştırılmasını engellemiştir. Ayrıca ilk 100 öznitelik içerisinde de başarı performansını düşüren özniteliklerin olabileceği düşünülmüş ve uygulanan 4 öznitelik seçim yöntemi içerisinde ilk 20 öznitelik ele alınarak en az 3 yöntem ile seçilen ilk 5 öznitelik belirlenmiştir.

Veri kümelerinde ortak seçilen öznitelikler Çizelge 4.5'te yer almaktadır. Ortak seçilen özniteliklerde donma tuş vuruş süresi öznitelikler işaretlenerek (boşluk tuş vuruşu, yeşil; sol tuş vuruşu, turuncu; sağ tuş vuruşu, mavi), Parkinson hastalığı klavye yazma davranışında boşluk tuşunun ve donma süresinin önemi vurgulanmıştır. Ayrıca donma süresi, bir tuşa basma ve bırakma arasında geçen süre olduğundan, Parkinson hastalarının kas kontrol mekanizmasında meydana gelen bozulmalardan dolayı katılık, donma ve hareketlerde yavaşlama belirtilerinin klavye yazma davranışı ile tespit edilebileceği görülmüştür.

Çizelge 4.5. Ortak seçilen ilk 5 öznitelik.

Ortak Seçilen Öznitelikler		1	2	3	4	5
1	Tüm_Sağlıklı (30) - Tüm_PH (88)	RS_gecikme Örnek Entropi	SS_donma Medyan	SS_donma Ortalama	RR_donma Aralık	RR_donma Maksimum
2	Tüm_Sağlıklı (30) - Tüm_PH (30 rastgele)	SS_donma Ortalama	RR_donma Aralık	SS_donma Maksimum	LR_gecikme HFBA	LS_donma Çarpıklık
3	Tüm_Sağlıklı(30) - Tüm_PH_ilacvar (78)	RS_gecikme Örnek Entropi	SS_donma Medyan	SS_donma Ortalama	RR_donma Maksimum	RR_donma Aralık
4	Tüm_Sağlıklı(30) - Tüm_PH_ilaçvar (30 rastgele)	SS_donma Ortalama	SS_donma aralık	RR_donma Maksimum	LS_donma Çarpıklık	RR_donma Aralık
5	Tüm_Sağlıklı(30) - Tüm_PH_ilaçyok (10)	SL_donma Maksimum	RS_donma Çarpıklık	LL_donma Standart Sapma	RS_donma Basıklık	SR_donma Baıklık
6	Tüm_Sağlıklı (30) - Hafif_PH (37)	RR_donma Aralık	RS_donma Çarpıklık	RR_donma Maksimum	LL_donma Standart Sapma	RS_donma Maksimum
7	Tüm_Sağlıklı (30) - Hafif_PH (10 rastgele)	RS_donma Basıklık	SR_donma Basıklık	LL_donma Örnek Entropi	LR_donma Basıklık	RS_donma Maksimum
8	Tüm_Sağlıklı (30) - Hafif_PH_ilaçvar (30)	RS_donma Maksimum	RR_donma Maksimum	RS_donma Çarpıklık	RR_donma Aralık	RR_donma Standart Sapma
9	Tüm_Sağlıklı (30) - Hafif_PH_ilaçyok (7)	LR_donma Aralık	RS_donma Maksimum	RS_donma Aralık	SL_donma Maksimum	RS_donma Çarpıklık
10	Tüm_Sağlıklı (30) - Orta_PH (39)	RS_gecikme Örnek Entropi	SS_donma Medyan	SS_donma Ortalama	SR_uçuş HFBA	SS_donma Maksimum
11	Tüm_Sağlıklı (30) - Orta_PH (10 rastgele)	RL_uçuş HFBA	SS_donma Ortalama	LS_gecikme HFBA	SL_donma Standart Sapma	LS_donma Çarpıklık
12	Tüm_Sağlıklı (30) - Ciddi_PH (10)	LR_donma Basıklık	LR_donma Çarpıklık	RS_donma Aralık	LR_gecikme HFBA	LR_donma Aralık
13	Kadın_Sağlıklı (16) - Kadın_PH (51)	SS_donma Ortalama	LS_donma Çarpıklık	SS_donma Medyan	SS_donma Satandart Sapma	SS_donma Aralık
14	Erkek_Sağlıklı (14) - Erkek_PH (37)	RS_gecikme Örnek Entropi	SS_donma Ortalama	RR_donma Maksimum	SR_uçuş HFBA	SS_donma Maksimum

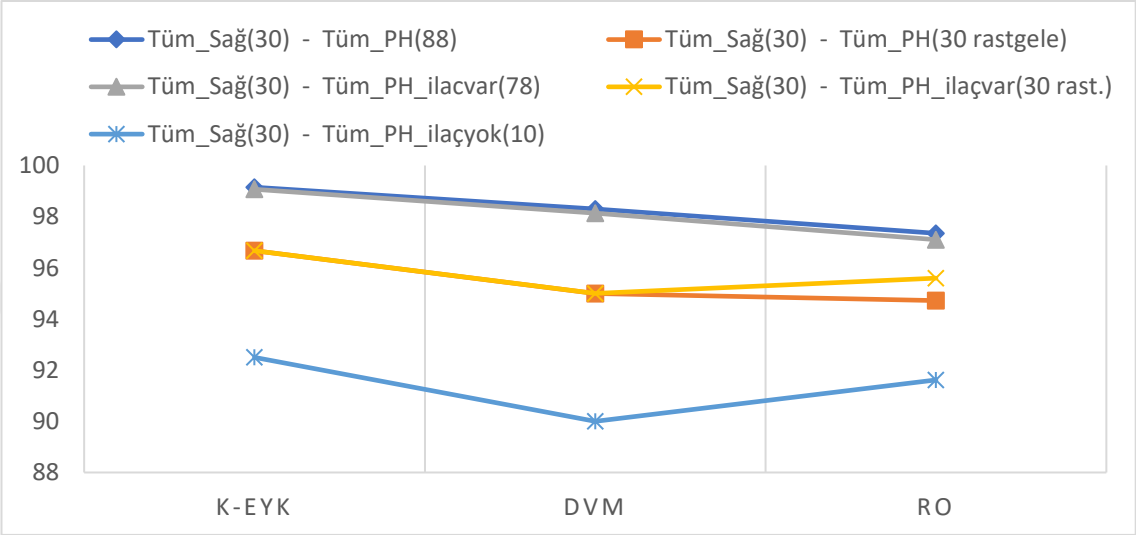
Veri kümeleri, ilk iki aşamada olduğu gibi ilaç kullanım durumları, hastalık şiddeti ve cinsiyete göre sınıflandırma algoritmalarının başarı performansları incelenmiştir. Bu aşamada ek olarak f-ölçütü ve MKK değerleri de incelenmiş ve değerlendirilmiştir. İstatiksel ölçüm yöntemlerinden olan f-ölçütü ve MKK değerlerinin 1'e yakın olması elde edilen sınıflandırma sonuçlarının başarısını göstermektedir. Çizelge 4.6 veri kümelerine ait doğruluk, f-ölçütü ve MKK değerlerini içermektedir.

Çizelge 4.6. Orta seçilen özniteliklere göre sınıflandırma başarı performansları.

Ortak Seçilen Öznitelikler		k-EYK			DVM			RO		
		Doğruluk	f-ölçütü	MKK	Doğruluk	f-ölçütü	MKK	Doğruluk	f-ölçütü	MKK
1	**Tüm_Sağlıklı (30) - Tüm_PH (88)**	**99,15**	0,99	0,97	98,30	0,98	0,95	97,34	0,97	0,89
2	**Tüm_Sağlıklı (30) - Tüm_PH (30 rastgele)**	**96,67**	0,96	0,93	95	0,94	0,90	94,70	0,93	0,86
3	**Tüm_Sağlıklı(30) - Tüm_PH_ilacvar (78)**	**99,07**	0,99	0,97	98,14	0,98	0,95	97,12	0,97	0,89
4	**Tüm_Sağlıklı(30) - Tüm_PH_ilaçvar (30 rastgele)**	**96,67**	0,96	0,93	95	0,94	0,90	95,67	0,93	0,86
5	**Tüm_Sağlıklı(30) - Tüm_PH_ilaçyok (10)**	**92,50**	0,94	0,93	90	0,83	0,79	91,50	0,75	0,69
6	**Tüm_Sağlıklı (30) - Hafif_PH (37)**	**97,01**	0,97	0,93	95,52	0,96	0,91	95,18	0,93	0,85
7	**Tüm_Sağlıklı (30) - Hafif_PH (10 rastgele)**	**95**	0,90	0,86	**95**	0,90	0,87	91,73	0,76	0,70
8	**Tüm_Sağlıklı (30) - Hafif_PH_ilaçvar (30)**	**96,67**	0,96	0,93	**96,67**	0,96	0,93	94,33	0,92	0,84
9	**Tüm_Sağlıklı (30) - Hafif_PH_ilaçyok (7)**	**91,89**	0,85	0,82	89,18	0,75	0,71	91,65	0,68	0,63
10	**Tüm_Sağlıklı (30) - Orta_PH (39)**	**95,65**	0,96	0,91	**95,65**	0,96	0,91	95,61	0,94	0,87
11	**Tüm_Sağlıklı (30) - Orta_PH (10 rastgele)**	**100**	1,00	1,00	97,50	0,94	0,92	93,53	0,82	0,77
12	**Tüm_Sağlıklı (30) - Ciddi_PH (10)**	**97,50**	0,95	0,93	**97,50**	0,95	0,94	95,70	0,74	0,69
13	**Kadın_Sağlıklı (16) - Kadın_PH (51)**	97,01	0,98	0,91	**98,50**	0,99	0,95	98,13	0,96	0,84
14	**Erkek_Sağlıklı (14) - Erkek_PH (37)**	**96,08**	0,97	0,90	96,07	0,97	0,90	95,63	0,95	0,83

Veri kümesi 1 ve 2'ye göre kişi sayılarının dengelenmesi sınıflandırma sonuçlarında doğruluk değerlerinin düşmesine neden olmuştur. En iyi sonuçlar k-EYK algoritmasında (%96,67 ve %99,15 doğruluk) k=5,9 komşuluğu ve City-Block mesafe ölçümünde elde edilirken, istatiksel ölçüm yöntemlerinden olan f-ölçütü ve MKK, 0,93 - 0,99 arasında değerler almıştır. İlaç kullanan hasta veri kümelerinde (veri kümesi 3, 4) de

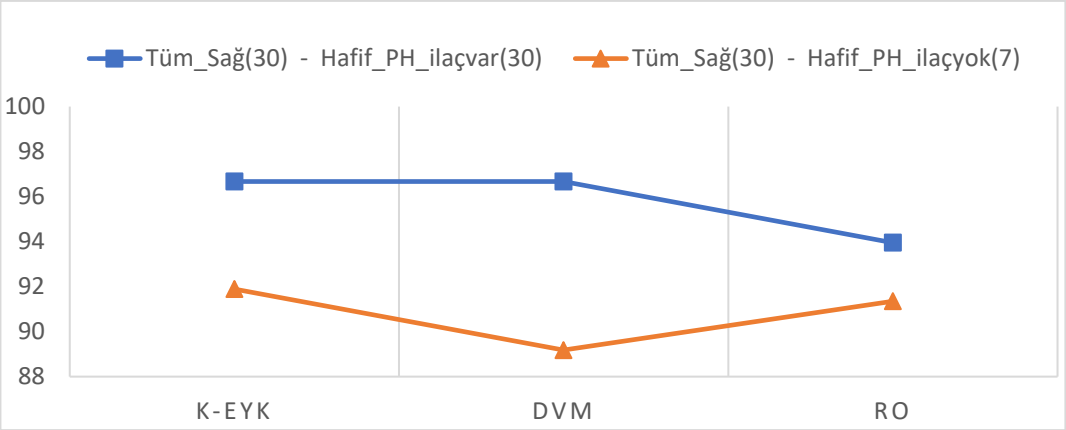
kişi sayısının dengeli hale getirilmesiyle doğruluk değerlerinde düşüşler tespit edilmiştir. k-EYK algoritması ile elde edilen sonuçlarda k=9 komşuluğu ve City-Block mesafe ölçümünün en yüksek doğruluk değerlerini veren parametreler olduğu görülmüştür. Şekil 4.9'da Tüm Parkinson hastaları ile oluşturulan veri kümelerinin doğruluk değerlerini temsil eden grafik verilmiştir.



Şekil 4.9. Tüm Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları.

Tüm Parkinson hastalarından meydana gelen veri kümelerinde ilaç kullanmayan katılımcıların olduğu veri kümesinin daha yüksek başarı performansı göstermesi beklenirken tam tersi sonuçlar elde edilmiştir. Bu durumun, veri kümeleri içerindeki katılımcıların hastalık şiddetlerinin bilinmemesinden kaynaklanabileceği düşünülerek hafif derece Parkinson hastalarında (veri kümesi 8, 9) da ilaç kullanım durumları incelenmiştir. İlaç kullanan hafif derece Parkinson hastalarında k-EYK algoritması ile %96,67 doğruluk değerine tüm komşuluklarda Chebychev mesafe ölçümüyle; ilaç kullanmayan hafif derece hastalarda ise %91,89 doğruluk değerine k=3 komşuluğunda City-Block mesafe ölçümüyle ulaşmıştır. f-ölçütü ve MKK değerleri ise 0,82-0,96 arasında değerler almıştır (Şekil 4.10).

Hafif derece Parkinson hastaları belirti göstermedikleri takdirde ilaç kullanmamaktadırlar. İlaç kullanan veri kümesinin daha yüksek sonuçlar vermesinin bir nedeni olarak, kullanılan ilaçlara rağmen hastaların daha önceden gösterdiği belirtilerin, klavye tuş vuruşları aracılığıyla tespit edebilmesi gösterilebilir.



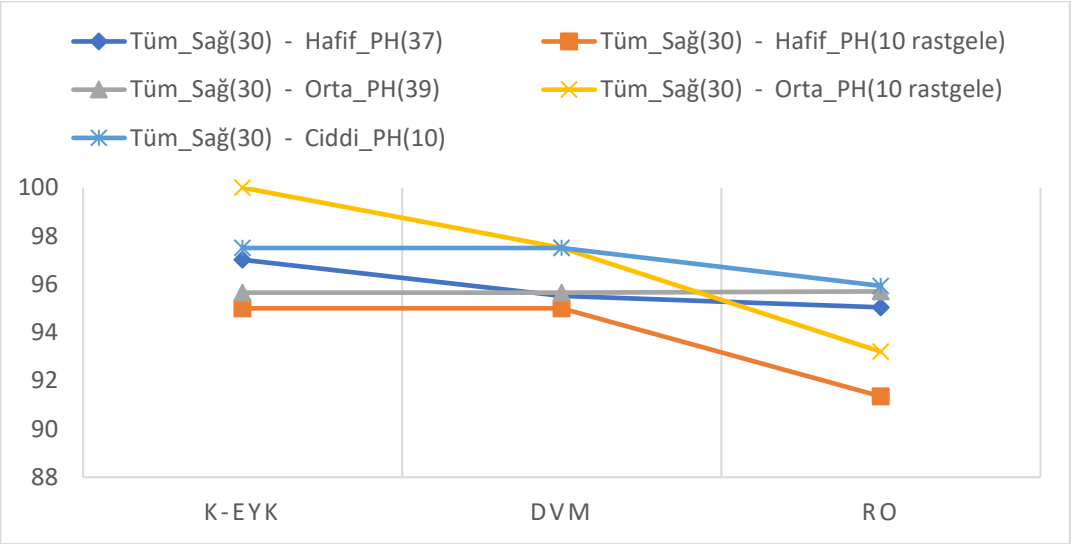
Şekil 4.10. Hafif derece Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları.

İlaç kullanım durumlarının ardından, Parkinson hastalığının şiddetine göre farklı sınıflandırıcılardan elde edilen sonuçlar incelenmiştir. Ciddi Parkinson hastaları ile oluşturulan veri kümesindeki katılımcı sayısı, diğer veri kümelerine göre daha az olduğundan dolayı, hastalık şiddetinde modelin performansını incelemek amacıyla hafif ve orta derece Parkinson hastalarından oluşan veri kümelerindeki katılımcılar rastgele seçilerek alt veri kümeleri oluşturulmuştur (veri kümesi 6, 7, 10-12).

K-EYK ve DVM algoritmaları hastalık şiddetinin kıyaslanmasında yakın sonuçlar vermiştir (Şekil 4.11). Katılımcı sayısının dengeli hale getirilmesiyle hafif, orta ve ciddi derece Parkinson hastaları k-EYK algoritmasında, sırasıyla %95, %100 ve %97,5 doğruluk değerlerini etmiş ve hepsinde ortak olarak k=3 komşuluğu ve Euclid ile Minkowski mesafe ölçüm parametrelerinin en başarılı sonuçları verdiği belirlenmiştir. DVM algoritmasında ise sırasıyla %95, %97,5 ve %97,5 doğruluk değerleri, t=2 (RBF) çekirdek fonksiyonu parametresi ile elde edilmiştir. F-ölçütü ve MKK değerleri ise 0,86-1,00 arasında değerler alarak modelin başarısını göstermektedir.

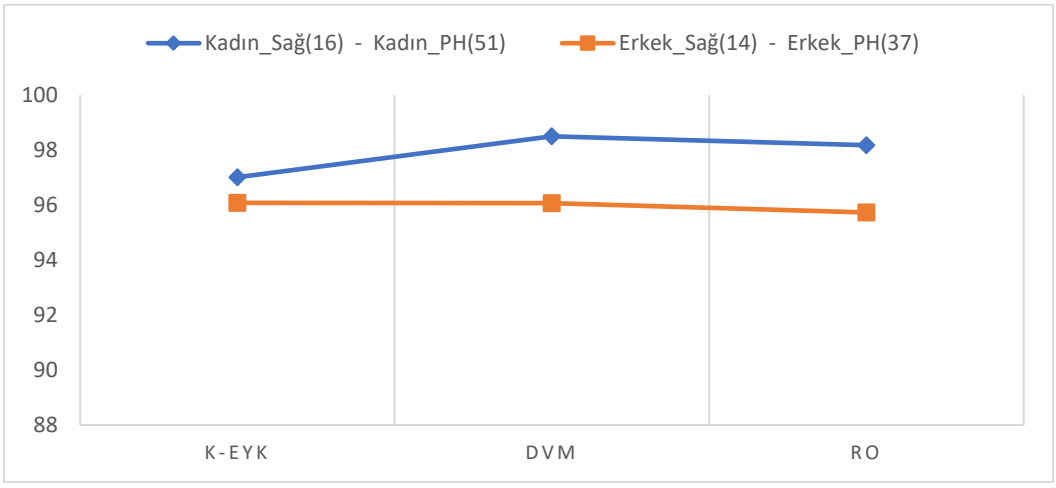
Hastalık şiddetinin artmasıyla, 378 öznitelik ve seçilen ilk 100 öznitelik ile elde edilen sınıflandırma sonuçlarında artış gözlemlenmemişti. Ortak seçilen özniteliklerle yapılan sınıflandırma işleminde de aynı durum oluşmuştur. Hastalığın şiddetinin artmasıyla kullanılan ilaç sayısı ve dozu artmaktadır. Ciddi Parkinson hastalarından oluşan veri kümesi incelendiğinde tüm hastaların en az bir adet ilaç kullandığı belirlenmiştir. Hafif ve orta derecede olan veri kümelerinde ise belirti göstermediği için

ilaç kullanmayan hastaların da bulunması sınıflandırma sonuçlarını beklenilenin tersi yönde olmasına neden olmuştur.



Şekil 4.11. Hastalık şiddetine göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları.

Son olarak cinsiyete göre yapılan sınıflandırma sonuçlarında (veri kümesi 13, 14), model kadın katılımcılarda daha başarılı hasta-sağlıklı ayrımı yapmıştır (Şekil 4.12). Kadın katılımcılar, %98,5 doğruluk değerini DVM algoritmasında, t=1 (polinom) parametresi ile elde ederken, erkek katılımcılarda ise k-EYK algoritması, k=5, 7, 9 komşuluk değerleri ve City-Block mesafe ölçümünde %96,08 başarı performansı göstermiştir.



Şekil 4.12. Cinsiyete göre Parkinson hastaları ve sağlıklı katılımcılardan oluşan veri kümelerinde ortak seçilen özniteliklerin başarı performansları.

Katılımcı sağlık bilgileri dikkate alınarak oluşturulan farklı veri kümeleri, normalizasyon yöntemlerine, öznitelik seçim yöntemlerine ve ortak seçilen özniteliklere göre değerlendirilmiştir. İlk iki aşamada modelin %100 doğruluk değerini vermesi aşırı öğrenmeden şüphelenilmesine neden olmuştur. Ayrıca f-ölçütü ve MKK gibi değerlendirme ölçütlerinin karşılaştırılamamasından dolayı, ortak seçilen öznitelikler aynı sınıflandırıcılarda çalıştırılarak doğruluk, f-ölçütü ve MKK sonuçları değerlendirilmiştir.

Ortak seçilen özniteliklerde boşluk tuş vuruşuna ve donma süresine ait özniteliklerin seçildiği görülmüştü. Elde edilen sonuçlara göre klavyede günlük yazma davranışı, boşluk tuşuna yapılan vuruşlarla Parkinson hastalığını yüksek başarım ile tespit edebilmektedir. Ayrıca donma, katılık ve hareketlerde yavaşlama belirtileri de tuş vuruşu süresini etkileyerek, Parkinson hastalığının tespitinde önemli olmuştur.

Çizelge 4.7'de öznitelik seçimi uygulanmadan olmadan önce sınıflandırma algoritmalarının performansları, öznitelik seçim yöntemleri uygulandıktan sonra en yüksek başarımı gösteren modellerin sonuçları ve ortak seçilen öznitelikler ile elde edilen sonuçlar karşılaştırılmaktadır.

Çizelge 4.7. Veri kümelerine ait sınıflandırma doğruluk değerleri.

	Veri Kümeleri	378 öznitelik			Öznitelik seçimi			Ortak 5 öznitelik		
		k-EYK	DVM	RO	k-EYK	DVM	RO	k-EYK	DVM	RO
1	Tüm_Sağlıklı(30)-Tüm_Parkinson(88)	**95.76**	**100**	**99.93**	**100** mRmR, Relief, SFS	**100** mRmR, Relief, SFS	**99.82** SFS	**99.15**	**98.3**	**97.35**
2	Tüm_Sağlıklı(30)-Tüm_Parkinson(30 rastgele)	**95**	**98.33**	**99.98**	**100** mRmR	**100** Relief	**99.33** SFS	**96.67**	**95**	**94.72**
3	Tüm_Sağlıklı(30)-Tüm_Parkinson_ilacvar(78)	**95.37**	**100**	**99.95**	**100** mRmR, Relief, SFS	**100** mRmR, Relief	**98.93** SFS	**99.07**	**98.14**	**97.1**
4	Tüm_Sağlıklı(30)-Tüm_Parkinson_ilacvar(30 rastgele)	**96.67**	**100**	**99.96**	**100** mRmR	**100** mRmR	**98.95** SFS	**96.67**	**95**	**95.6**
5	Tüm_Sağlıklı(30)-Tüm_Parkinson_ilacyok(10)	**82.5**	**95**	**98.6**	**95** mRmR, SFS	**100** mRmR, RO	**94.88** SFS	**92.5**	**90**	**91.62**
6	Tüm_Sağlıklı(30)-Parkinson_hafif(37)	**98.5**	**100**	**99.91**	**100** Relief	**100** mRmR, Relief	**98.84** mRmR, Relief	**97.01**	**95.52**	**95.04**
7	Tüm_Sağlıklı(30)-Parkinson_hafif(10 rastgele)	**87.5**	**92.5**	**99.92**	**100** mRmR	**97.5** mRmR, Relief,	**94.6** SFS	**95**	**95**	**91.35**
8	Tüm_Sağlıklı(30)-Parkinson_hafif_ilacvar(30)	**98.33**	**98.33**	**99.92**	**100** mRmR, Relief,	**100** mRmR, Relief	**98.85** SFS	**96.67**	**96.67**	**93.95**
9	Tüm_Sağlıklı(30)-Parkinson_hafif_ilacyok(7)	**81.08**	**97.29**	**98.24**	**97.9** mRmR	**100** mRmR	**93.57** RO	**91.89**	**89.18**	**91.35**
10	Tüm_Sağlıklı(30)-Parkinson_orta(39)	**94.2**	**100**	**99.75**	**100** mRmR, SFS	**98.55** Relief	**99.33** mRmR, Relief	**95.65**	**95.65**	**95.7**
11	Tüm_Sağlıklı(30)-Parkinson_orta(10 rastgele)	**82.5**	**90**	**99.95**	**97.5** RO	**100** mRmR, Relief	**96.62** RO	**100**	**97.5**	**93.2**
12	Tüm_Sağlıklı(30)-Parkinson_ciddi(10)	**87.5**	**90**	**99.52**	**100** mRmR	**97.5** mRmR, Relief, SFS	**96.4** SFS	**97.5**	**97.5**	**95.93**
13	kadın_Sağlıklı(16)-kadın_Parkinson(51)	**88.59**	**98.5**	**99.96**	**100** mRmR	**100** mRmR	**98.57** SFS	**97.01**	**98.5**	**98.18**
14	erkek_Sağlıklı(14)-erkek_Parkinson(37)	**88.23**	**98.03**	**98.92**	**100** mRmR	**100** mRmR	**96.1** SFS	**96.08**	**96.07**	**95.73**

BÖLÜM 5

SONUÇ VE TARTIŞMA

Parkinson hastalığı sinir hasarlarından meydana gelen bir rahatsızlık olmasından dolayı hastalara uygulanabilecek kesin bir tedavi yöntemi bulunmamaktadır. Yapılan ilaç tedavileri, sadece hastanın günlük yaşamını kolaylaştırmak ve hastalığın ilerlemesini engellemeyi amaçlamaktadır. Bu sebeple hastalığın en erken evrede belirlenmesi hasta ve hasta yakınları açısından büyük önem taşımaktadır. Hastalığın ilk evrelerinde belirtiler genellikle el ve kollarda oluştuğundan, literatürde yapılan çalışmalar incelendiğinde, parmak vuruş hareketlerinin hastalığın erken evre tespitinde oldukça başarılı olan yöntemlerden biri olduğu gözlemlenmiştir. Parmak vuruş hareketlerinde, işaret ve baş parmağın birbirine vurulması veya klavyede yazma davranışı ile elde edilen süreler baz alınarak yapılan çalışmalarda hastalığın erken evrede tespitinin mümkün olduğu belirlenmiştir.

Bu tez çalışmasında, Parkinson hastalarına ait günlük klavye tuş vuruş verilerinden elde edilen öznitelikler kullanılarak hasta-sağlıklı ayrımı yapan bir model oluşturulmuş ve bu modelin uzmanlara hastalığın teşhisinde yardımcı olması amaçlanmıştır.

Physionet veri tabanından, Parkinson hastalarına ait günlük klavye tuş vuruş veri kümesi alınmış, hastalara ait sağlık bilgileri kullanılarak veri kümesi uygun alt veri kümelerine ayrılmıştır. Her veri kümesi ayrı ayrı incelenerek hastalık derecesine, ilaç kullanım durumlarına ve cinsiyete göre veri kümeleri arasında karşılaştırmalar yapılmıştır. Veride klavyede yapılan tuş vuruşlarına ait donma, gecikme ve uçuş

sürelerine ait yer alan bilgiler kullanılarak Matlab yazılım platformunda öznitelikler çıkarılmış ve hasta-sağlıklı ayrımı yapabilen bir eğitim modeli geliştirilmiştir. Toplamda 378 öznitelik çıkarılmış ve bu özniteliklerin önemi araştırılmıştır.

Öznitelik seçim yöntemleri ile yapılan sınıflandırmadan önce farklı normalizasyon yöntemleri uygulanmış ve hangisinin hasta sağlıklı ayrımında daha başarılı olduğu belirlenmiştir. Bu çalışmada, k-EYK ve DVM ile yapılan sınıflandırma sonuçlarına göre minimum-maksimum normalizasyon yöntemi neredeyse bütün veri kümelerinde en iyi performansa sahip olan yöntem olmuştur. RO algoritmasında minimum-maksimum normalizasyon ve z-skor normalizasyon arasında anlamlı bir farklılık görülmemiştir.

Seçilen öznitelikler ile yapılan sınıflandırmalara ait sonuçlar her veri kümesi için ayrı ayrı incelenmiş ve hangi yöntemin (mRmR, Relief, SFS, RO) daha başarı olduğu belirlenmiştir. Toplamda 378 öznitelik arasından belirlenen en başarılı ilk 100 öznitelik incelenmiş ve modelin en az sayıda öznitelik ile yüksek doğruluk değerlerine ulaşması amaçlanmıştır. Farklı sınıflandırma algoritmaları (k-EYK, DVM, RO) uygulanarak, sınıflandırma yöntemlerinin de veri kümeleri arasındaki başarıları karşılaştırılmıştır. mRmR yönteminin neredeyse tüm sınıflandırma sonuçlarında yüksek başarım sağladığı görülmüştür.

Farklı öznitelik seçim yöntemleri ile seçilen öznitelikler kullanılarak elde edilen sınıflandırma sonuçlarında birbirine yakın yüksek doğruluk sonuçları elde edilmiştir. Algoritmalar arasında karşılaştırma yapmayı engelleyen bu durumun önüne geçmek amacıyla en az 3 öznitelik seçim yöntemi ile seçilen ilk 5 öznitelik belirlenmiş ve sınıflandırma sonuçları elde edilmiştir. Ayrıca seçilen özniteliklerin büyük çoğunluğunu donma sürelerine ve boşluk tuş vuruşlarına ait olduğu gözlemlenmiştir. Veri kümesi içerisinde sol ya da sağ tuşlara yapılan vuruşların hangi tuşa yapıldığı bilinmemesine rağmen boşluk tuşuna yapılan vuruşlar bilinmekteydi. Bu nedenle, model Parkinson hastalığı tespitinde boşluk tuş vuruşlarının seçilmesiyle daha yüksek başarı performansına ulaşmıştır. Donma süresi ise tek tuşa yapılan vuruştaki geçen süreyi temsil ettiğinden dolayı sağ-sol geçişlerden kaynaklanan negatif etkilerden etkilenmemektedir. Ayrıca donma süresinin tek tuşa bağlı olması, klavyede yazma hızının ve daha önce klavye kullanmanın önemini ortadan kaldırmaktadır. Ortak seçilen ilk 5 öznitelik

içerisinde donma süresinin yer alması yüksek doğruluk değerlerinin elde edilmesini sağlamıştır.

Hastalık şiddetine, ilaç kullanım durumuna, cinsiyete ve veri kümelerindeki katılımcı sayılarını dengelemek amacıyla oluşturulan veri kümelerinde elde edilen sonuçlara göre, tüm Parkinson hastaları ve tüm sağlıklılarla oluşturulan veri kümelerinde %90-100 arasında doğruluk değerleri elde edilmiştir. Katılımcı sayılarını dengelemek, doğruluk değerlerinde bazı veri kümelerinde artan bazılarında azalan bir değişim göstermiştir.

Tüm Parkinson ve tüm sağlıklı veri kümelerinde dengeli veri kümeleri oluşturmak doğruluk değerlerini çok etkilememiştir. Herhangi bir ilaç kullanan hastalar, ilaç kullanmayan hastalara göre daha yüksek doğruluk değerleri elde etmişlerdir. Veri kümesindeki katılımcıların hastalık şiddetinin bilinmemesinden dolayı sadece hafif derece Parkinson hastalarından oluşan alt veri kümeleri oluşturulmuştur. Hafif derece Parkinson hastaları ve sağlıklı bireylerle oluşturulan veri kümelerinden, %81.08-100 arasında doğruluk değerleri elde edilmiştir. İlaç kullanan hastalar, kullanmayan hastalara göre daha yüksek sınıflandırma sonuçları elde etmiştir. Bunun bir nedeni belirtilere sahip hafif Parkinson hastaları ilaç kullansalar dahi hasta bireylere göre, klavye tuş vuruşu esnasında milisaniye cinsinden farklılıklara sahip olmasıdır.

Hafif, orta ve ciddi Parkinson hastaları ve sağlıklılardan oluşan veri kümelerinde ciddi Parkinson hastalarında katılımcı sayısı az olduğundan, rastgele seçilen kişilerle farklı veri kümeleri oluşturulmuş ve karşılaştırılmıştır. Tüm sağlıklı katılımcıların ve şiddetine göre oluşturulan hasta veri kümelerinin sınıflandırma sonuçlarına göre, RO'da algoritmanın 100 defa çalıştırılıp ortalamasının alınması, hastalığın şiddeti arttıkça doğruluk değerinin artmasında etkili olmuştur.

Cinsiyet veri kümelerinde ise, kadın katılımcılardan oluşan Parkinson hastaları veri kümesi erkek katılımcılardan oluşan veri kümesine göre daha yüksek ancak yakın doğruluk değerleri elde etmiştir.

Klavye tuş vuruş dinamikleri, kişiye özgü vurma davranışlarını içermekte olup literatürde klavye ile yazma işlemi sırasında kimlik tespiti ve hasta-sağlıklı ayrımı yapılmaktadır. Parkinson hastaları erken evrelerinde, hareket yavaşlaması, titreme ve donma gibi belirtiler mevcuttur. El ve kollarda meydana gelen ilk belirtiler hastalığın

ilerlemesiyle baş, ayak ve seste de ortaya çıkmaktadır. Bu çalışmada klavyede yapılan parmak tuş vuruş süreleri incelenerek boşluk tuş vuruşunun, donma süresinin Parkinson hastalığını ayırt edici özellikte olduğu belirlenmiştir.

Yapılan tez çalışmasında, klavye tuş vuruş dinamikleri incelenerek katılımcının sağlık bilgileri doğrultusunda veri kümeleri oluşturulmuş ve Parkinson hastalarını sağlıklı katılımcılardan ayıran bir model geliştirilmiştir. İleriki çalışmalarda, Leap Motion Controller cihazı ile toplanan parmak vuruşu, el açma kapama ve el döndürme hareketi verileri yardımıyla hasta-sağlıklı ayrımı ve hastalık evre tespiti yapan, makine öğrenimi yöntemlerine ek olarak derin öğrenme yöntem ve tekniklerinin de kullanıldığı hiyerarşik bir modelin kurulması hedeflenmiştir. Ayrıca oluşturulan modelin, sonraki çalışmalarda sadece Parkinson hastalığı değil sinir sisteminin etkilenmesiyle oluşan diğer hastalıkların (Huntington, Multipl skleroz, Otizm vb.) tespitinde ve takibinde de kullanılması amaçlanmaktadır.

KAYNAKLAR

Adams, W. R. (2017). High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing. *PloS one*, *12*(11), e0188226.

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

Akar, S. A., Kara, S., Agambayev, S., & Bilgiç, V. (2015, October). Investigation of EEG signals of patients with major depression using chaotic features. In 2015 Medical Technologies National Conference (TIPTEKNO) (pp. 1-4). IEEE.

Akpınar, H. (2017). Data Veri Madenciliği - Veri Analizi: Papatya

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.

Arroyo-Gallego, T., Ledesma-Carbayo, M. J., Sánchez-Ferro, A., Butterworth, I., Mendoza, C. S., Matarazzo, M., ... & Giancardo, L. (2017). Detection of motor impairment in Parkinson's disease via mobile touchscreen typing. *IEEE Transactions on Biomedical Engineering*, *64*(9), 1994-2002.

Ashour, A. S., Nour, M. K. A., Polat, K., Guo, Y., Alsaggaf, W., & El-Attar, A. (2020). A novel framework of two successive feature selection levels using weight-based procedure for voice-loss detection in Parkinson's disease. IEEE Access, 8, 76193-76203.

Asmae, O., Abdelhadi, R., Bouchaib, C., Sara, S., & Tajeddine, K. (2020, April). Parkinson's disease identification using KNN and ANN Algorithms based on Voice Disorder. In 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET) (pp. 1-6). IEEE.

Atasoy, H., Yıldırım, S., Yıldırım, E., & Kutlu, Y. (2014). EEG Sinyallerinden Fraktal Boyut Ve Dalgacık Dönüşümü Kullanılarak Duygu Tanıma Emotion Recognition from EEG Signals Using Fractal Dimension And Wavelet Transform.

Ayodele, T. O. (2010). Types of machine learning algorithms. New advances in machine learning, 3, 19-48. https://books.google.com.tr/books

Bhidayasiri, R., & Tarsy, D. (2012). *Movement disorders: a video atlas*. Springer Science & Business Media. (s. 4-5). https://books.google.com.tr/books

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Butt, A. H., Rovini, E., Dolciotti, C., De Petris, G., Bongioanni, P., Carboncini, M. C., & Cavallo, F. (2018). Objective and automatic classification of Parkinson disease with Leap Motion controller. *Biomedical engineering online*, *17*(1), 1-21.

Chen, X., Solomon, I. C., & Chon, K. H. (2006, January). Comparison of the use of approximate entropy and sample entropy: applications to neural respiratory signal. In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference (pp. 4212-4215). IEEE.

Choi, J. H., Ma, H. I., Kim, Y. J., Lee, U., Bhatia, P. K., & Sharma, A. (2016). Development of an assessment method of forearm pronation/supination motor function based on mobile phone accelerometer data for an early diagnosis of Parkinson's disease. *Int J Bio-Science Bio-Technology*, *8*(2), 1-10.

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.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory, 13*(1), 21-27.

Çalışkan, S. K., & Soğukpınar, İ. (2008). KxKNN: K-Means ve K En Yakin Komşu Yöntemleri İle Ağlarda Nüfuz Tespiti. EMO Yayınları, 120-24.

Ding, C., & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology*, *3*(02), 185-205.

ER, M. B. (2021). Özellik seçimi ve Derin Öğrenmeye Dayalı Parkinson Hastalığı Tespiti. *Avrupa Bilim ve Teknoloji Dergisi*, *(21)*, 428-436.

Gao, C., Smith, S., Lones, M., Jamieson, S., Alty, J., Cosgrove, J., ... & Chen, S. (2018). Objective assessment of bradykinesia in Parkinson's disease using evolutionary algorithms: clinical validation. *Translational neurodegeneration*, *7*(1), 1-8.

Garner, D. M., De Souza, N. M., & Vanderlei, L. C. M. (2018). Heart rate variability analysis: Higuchi and Katz's fractal dimensions in subjects with type 1 diabetes mellitus. *Romanian Journal of Diabetes Nutrition and Metabolic Diseases*, *25*(3), 289-295.

Giancardo, L., Sánchez-Ferro, A., Butterworth, I., Mendoza, C. S., & Hooker, J. M. (2015). Psychomotor impairment detection via finger interactions with a computer keyboard during natural typing. *Scientific reports*, *5*(1), 1-8.

Giancardo, L., Sanchez-Ferro, A., Arroyo-Gallego, T., Butterworth, I., Mendoza, C. S., Montero, P., ... & Estépar, R. S. J. (2016). Computer keyboard interaction as an indicator of early Parkinson's disease. *Scientific reports*, *6*(1), 1-10.

Goetz, C. G. (2011). The history of Parkinson's disease: early clinical descriptions and neurological therapies. *Cold Spring Harbor perspectives in medicine*, *1*(1), a008862.

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *circulation*, *101*(23), e215-e220.

Gómez, P., Mekyska, J., Gómez, A., Palacios, D., Rodellar, V., & Álvarez, A. (2019). Characterization of Parkinson's disease dysarthria in terms of speech articulation kinematics. Biomedical Signal Processing and Control, 52, 312-320.

Gülgezen, G., Çataltepe, Z., & Yu, L. (2009). MRMR Algoritması Kullanılarak Kararlı Öznitelik Seçimi Stable Feature Selection Using MRMR Algorithm.

https://www.elcerrahi.com/basit-el-anatomisi.html

https://www.hemensaglik.com/makale/evinizi-parkinson-hastaligina-uygun-hale-getirmek

https://www.kaggle.com/residentmario/kernels-and-support-vector-machine-regularization

Jones, L. A., & Lederman, S. J. (2006). *Human hand function*. Oxford university press. (14-18). https://books.google.com.tr/books

Kandori, A., Yokoe, M., Sakoda, S., Abe, K., Miyashita, T., Oe, H., ... & Tsukada, K. (2004). Quantitative magnetic detection of finger movements in patients with Parkinson's disease. *Neuroscience research*, *49*(2), 253-260.

KAYNAR, O., Arslan, H., Görmez, Y., & IŞIK, Y. E. (2018). Makine öğrenmesi ve öznitelik seçim yöntemleriyle saldırı tespiti. *Bilişim Teknolojileri Dergisi, 11*(2), 175-185.

Kim, J. W., Lee, J. H., Kwon, Y., Kim, C. S., Eom, G. M., Koh, S. B., ... & Park, K. W. (2011). Quantification of bradykinesia during clinical finger taps using a gyrosensor in patients with Parkinson's disease. *Medical & biological engineering & computing*, *49*(3), 365-371.

Kira, K., & Rendell, L. A. (1992). A practical approach to feature selection. In Machine learning proceedings 1992 (pp. 249-256). Morgan Kaufmann.

Kurt, İ., Ulukaya, S., & Erdem, O. (2018, July). Musical feature based classification of parkinson's disease using dysphonic speech. In 2018 41st International Conference on Telecommunications and Signal Processing (TSP) (pp. 1-4). IEEE.

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

Lee, C. Y., Kang, S. J., Hong, S. K., Ma, H. I., Lee, U., & Kim, Y. J. (2016). A validation study of a smartphone-based finger tapping application for quantitative assessment of bradykinesia in Parkinson's disease. *PloS one*, *11*(7), e0158852.

Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on information theory*, *22*(1), 75-81.

Li, C., Zhang, X., & Jin, L. (2017). LPSNet: a novel log path signature feature based hand gesture recognition framework. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 631-639).

Liu, Y., Li, Y., Tan, X., Wang, P., & Zhang, Y. (2021). Local discriminant preservation projection embedded ensemble learning based dimensionality reduction of speech data of Parkinson's disease. Biomedical Signal Processing and Control, 63, 102165.

Lu, R., Xu, Y., Li, X., Fan, Y., Zeng, W., Tan, Y., ... & Cao, X. (2020). Evaluation of wearable sensor devices in Parkinson's disease: A review of current status and future prospects. Parkinson's Disease, 2020.

Manyam, B. V. (1990). Paralysis agitans and levodopa in “Ayurveda”: ancient Indian medical treatise. *Movement disorders: official journal of the Movement Disorder Society*, *5*(1), 47-48.

Marcano-Cedeño, A., Quintanilla-Domínguez, J., Cortina-Januchs, M. G., & Andina, D. (2010, November). Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network. In IECON 2010-36th annual conference on IEEE industrial electronics society (pp. 2845-2850). IEEE.

Moshkova, A., Samorodov, A., Voinova, N., Volkov, A., Ivanova, E., & Fedotova, E. (2020, April). Parkinson’s disease detection by using machine learning algorithms and hand movement signal from LeapMotion sensor. In 2020 26th Conference of Open Innovations Association (FRUCT) (pp. 321-327). IEEE.

Mounika, P., & Rao, S. G. (2021). A Study On Deep Learning And Machine Learning Techniques On Detection Of Parkinson’s Disease. In E3S Web of Conferences (Vol. 309). EDP Sciences.

Müller, B., Assmus, J., Herlofson, K., Larsen, J. P., & Tysnes, O. B. (2013). Importance of motor vs. non-motor symptoms for health-related quality of life in early Parkinson's disease. *Parkinsonism & related disorders*, *19*(11), 1027-1032.

Özekmekçi, S., Apaydın, H., Oğuz, S., & Zileli, İ. (2008). *Parkinson Hastalığı Hasta ve Yakınları İçin El Kitabı*: Parkinson Hastalığı Derneği

Peachap, A. B., Tchiotsop, D., Louis-Dorr, V., & Wolf, D. (2020). Detection of early Parkinson’s disease with wavelet features using finger typing movements on a keyboard. *SN Applied Sciences*, *2*(10), 1-8.

Perlmutter, J. S. (2009). Assessment of Parkinson disease manifestations. *Current protocols in neuroscience*, *49*(1), 10-1.

Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In Machine Learning (pp. 101-121). Academic Press.

Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., ... & Lang, A. E. (2017). Parkinson disease. *Nature reviews Disease primers*, *3*(1), 1-21.

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.

Rosas, S., Reynoso, A., Lopez, O., Cruz, I., Osorio, R., Díaz, I., ... & Rotenberg, N. K. (1747). Compartmental modeling of F-DOPA PET images from Parkinson’s patients. In *AIP Conference Proceedings* (Vol. 80012, No. 2016).

Schaap, T. (2018). Predicting Severity of Parkinson's Disease with Typing Behavior: A Machine Learning Approach (Master's thesis).

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379-423.

Solana-Lavalle, G., Galán-Hernández, J. C., & Rosas-Romero, R. (2020). Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features. *Biocybernetics and Biomedical Engineering*, *40*(1), 505-516.

Stern, G. E. R. A. L. D. (1989). Did parkinsonism occur before 1817?. *Journal of neurology, neurosurgery, and psychiatry*, *52*(Suppl), 11.

Tamson, N. (2018). *Parkinson Hastalığının Tuş Vuruş Datası Kullanılarak Tespiti*. (Yüksek Lisans Tezi). Başkent Üniversitesi/Fen Bilimleri Enstitüsü, Ankara.

VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data.* " O'Reilly Media, Inc.". (s. 332). https://books.google.com.tr/books

Vásquez-Correa, J. C., Arias-Vergara, T., Orozco-Arroyave, J. R., Eskofier, B., Klucken, J., & Nöth, E. (2018). Multimodal assessment of Parkinson's disease: a deep learning approach. *IEEE journal of biomedical and health informatics*, *23*(4), 1618-1630.

Xia, R., Muthumani, A., Mao, Z. H., & Powell, D. W. (2016). Quantification of neural reflex and muscular intrinsic contributions to parkinsonian rigidity. *Experimental brain research*, *234*(12), 3587-3595.

Xia, Y., Yao, Z., Ye, Q., & Cheng, N. (2019). A dual-modal attention-enhanced deep learning network for quantification of Parkinson's disease characteristics. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(1), 42-51.

Yılancıoğlu, K. (2017). Vocal Cord Measures Based Artificial Neural Network Approach for Prediction of Parkinson's Disease Status. *SDU Journal of Health Science Institute/SDÜ Saglik Bilimleri Enstitüsü Dergisi*, (1).

Zhang, H., Song, C., Rathore, A. S., Huang, M. C., Zhang, Y., & Xu, W. (2020). mHealth Technologies Towards Parkinson's Disease Detection and Monitoring in Daily Life: A Comprehensive Review. *IEEE reviews in biomedical engineering*, *14*, 71-81.

EKLER

EK 1. Klavye tuş vuruşlarından çıkarılan öznitelikler.

1	LL_donma_Minimum	26	LR_donma_ÖrnekEntropi
2	LL_donma_Maksimum	27	LR_donma_HFBA
3	LL_donma_Ortalama	28	LR_donma_KFBA
4	LL_donma_Medyan	29	LS_donma_Minimum
5	LL_donma_StandartSapma	30	LS_donma_Maksimum
6	LL_donma_Aralık	31	LS_donma_Ortalama
7	LL_donma_Entropi	32	LS_donma_Medyan
8	LL_donma_Çarpıklık	33	LS_donma_StandartSapma
9	LL_donma_Basıklık	34	LS_donma_Aralık
10	LL_donma_KolmogorovKarmaşıklığı	35	LS_donma_Entropi
11	LL_donma_LempelZivKarmaşıklığı	36	LS_donma_Çarpıklık
12	LL_donma_ÖrnekEntropi	37	LS_donma_Basıklık
13	LL_donma_HFBA	38	LS_donma_KolmogorovKarmaşıklığı
14	LL_donma_KFBA	39	LS_donma_LempelZivKarmaşıklığı
15	LR_donma_Minimum	40	LS_donma_ÖrnekEntropi
16	LR_donma_Maksimum	26	LR_donma_ÖrnekEntropi
17	LR_donma_Ortalama	41	LS_donma_HFBA
18	LR_donma_Medyan	42	LS_donma_KFBA
19	LR_donma_StandartSapma	43	RR_donma_Minimum
20	LR_donma_Aralık	44	RR_donma_Maksimum
21	LR_donma_Entropi	45	RR_donma_Ortalama
22	LR_donma_Çarpıklık	46	RR_donma_Medyan
23	LR_donma_Basıklık	47	RR_donma_StandartSapma
24	LR_donma_KolmogorovKarmaşıklığı	48	RR_donma_Aralık

EK 1. Klavye tuş vuruşlarından çıkarılan öznitelikler (devam).

49	RR_donma_Entropi	79	RS_donma_Basıklık
50	RR_donma_Çarpıklık	80	RS_donma_KolmogorovKarmaşıklığı
51	RR_donma_Basıklık	81	RS_donma_LempelZivKarmaşıklığı
52	RR_donma_KolmogorovKarmaşıklığı	82	RS_donma_ÖrnekEntropi
53	RR_donma_LempelZivKarmaşıklığı	83	RS_donma_HFBA
54	RR_donma_ÖrnekEntropi	84	RS_donma_KFBA
55	RR_donma_HFBA	85	SS_donma_Minimum
56	RR_donma_KFBA	86	SS_donma_Maksimum
57	RL_donma_Minimum	87	SS_donma_Ortalama
58	RL_donma_Maksimum	88	SS_donma_Medyan
59	RL_donma_Ortalama	89	SS_donma_StandartSapma
60	RL_donma_Medyan	90	SS_donma_Aralık
61	RL_donma_StandartSapma	91	SS_donma_Entropi
62	RL_donma_Aralık	92	SS_donma_Çarpıklık
63	RL_donma_Entropi	93	SS_donma_Basıklık
64	RL_donma_Çarpıklık	94	SS_donma_KolmogorovKarmaşıklığı
65	RL_donma_Basıklık	95	SS_donma_LempelZivKarmaşıklığı
66	RL_donma_KolmogorovKarmaşıklığı	96	SS_donma_ÖrnekEntropi
67	RL_donma_LempelZivKarmaşıklığı	97	SS_donma_HFBA
68	RL_donma_ÖrnekEntropi	98	SS_donma_KFBA
69	RL_donma_HFBA	99	SL_donma_Minimum
70	RL_donma_KFBA	100	SL_donma_Maksimum
71	RS_donma_Minimum	101	SL_donma_Ortalama
72	RS_donma_Maksimum	102	SL_donma_Medyan
73	RS_donma_Ortalama	103	SL_donma_StandartSapma
74	RS_donma_Medyan	104	SL_donma_Aralık
75	RS_donma_StandartSapma	105	SL_donma_Entropi
76	RS_donma_Aralık	106	SL_donma_Çarpıklık
77	RS_donma_Entropi	107	SL_donma_Basıklık
78	RS_donma_Çarpıklık	108	SL_donma_KolmogorovKarmaşıklığı

EK 1. Klavye tuş vuruşlarından çıkarılan öznitelikler (devam).

109	SL_donma_LempelZivKarmaşıklığı	139	LL_gecikme_HFBA
110	SL_donma_ÖrnekEntropi	140	LL_gecikme_KFBA
111	SL_donma_HFBA	141	LR_gecikme_Minimum
112	SL_donma_KFBA	142	LR_gecikme_Maksimum
113	SR_donma_Minimum	143	LR_gecikme_Ortalama
114	SR_donma_Maksimum	144	LR_gecikme_Medyan
115	SR_donma_Ortalama	145	LR_gecikme_StandartSapma
116	SR_donma_Medyan	146	LR_gecikme_Aralık
117	SR_donma_StandartSapma	147	LR_gecikme_Entropi
118	SR_donma_Aralık	148	LR_gecikme_Çarpıklık
119	SR_donma_Entropi	149	LR_gecikme_Basıklık
120	SR_donma_Çarpıklık	150	LR_gecikme_KolmogorovKarmaşıklığı
121	SR_donma_Basıklık	151	LR_gecikme_LempelZivKarmaşıklığı
122	SR_donma_KolmogorovKarmaşıklığı	152	LR_gecikme_ÖrnekEntropi
123	SR_donma_LempelZivKarmaşıklığı	153	LR_gecikme_HFBA
124	SR_donma_ÖrnekEntropi	154	LR_gecikme_KFBA
125	SR_donma_HFBA	155	LS_gecikme_Minimum
126	SR_donma_KFBA	156	LS_gecikme_Maksimum
127	LL_gecikme_Minimum	157	LS_gecikme_Ortalama
128	LL_gecikme_Maksimum	158	LS_gecikme_Medyan
129	LL_gecikme_Ortalama	159	LS_gecikme_StandartSapma
130	LL_gecikme_Medyan	160	LS_gecikme_Aralık
131	LL_gecikme_StandartSapma	161	LS_gecikme_Entropi
132	LL_gecikme_Aralık	162	LS_gecikme_Çarpıklık
133	LL_gecikme_Entropi	163	LS_gecikme_Basıklık
134	LL_gecikme_Çarpıklık	164	LS_gecikme_KolmogorovKarmaşıklığı
135	LL_gecikme_Basıklık	165	LS_gecikme_LempelZivKarmaşıklığı
136	LL_gecikme_KolmogorovKarmaşıklığı	166	LS_gecikme_ÖrnekEntropi
137	LL_gecikme_LempelZivKarmaşıklığı	167	LS_gecikme_HFBA
138	LL_gecikme_ÖrnekEntropi	168	LS_gecikme_KFBA

EK 1. Klavye tuş vuruşlarından çıkarılan öznitelikler (devam).

169	RR_gecikme_Minimum	199	RS_gecikme_Ortalama
170	RR_gecikme_Maksimum	200	RS_gecikme_Medyan
171	RR_gecikme_Ortalama	201	RS_gecikme_StandartSapma
172	RR_gecikme_Medyan	202	RS_gecikme_Aralık
173	RR_gecikme_StandartSapma	203	RS_gecikme_Entropi
174	RR_gecikme_Aralık	204	RS_gecikme_Çarpıklık
175	RR_gecikme_Entropi	205	RS_gecikme_Basıklık
176	RR_gecikme_Çarpıklık	206	RS_gecikme_KolmogorovKarmaşıklığı
177	RR_gecikme_Basıklık	207	RS_gecikme_LempelZivKarmaşıklığı
178	RR_gecikme_KolmogorovKarmaşıklığı	208	RS_gecikme_ÖrnekEntropi
179	RR_gecikme_LempelZivKarmaşıklığı	209	RS_gecikme_HFBA
180	RR_gecikme_ÖrnekEntropi	210	RS_gecikme_KFBA
181	RR_gecikme_HFBA	211	SS_gecikme_Minimum
182	RR_gecikme_KFBA	212	SS_gecikme_Maksimum
183	RL_gecikme_Minimum	213	SS_gecikme_Ortalama
184	RL_gecikme_Maksimum	214	SS_gecikme_Medyan
185	RL_gecikme_Ortalama	215	SS_gecikme_StandartSapma
186	RL_gecikme_Medyan	216	SS_gecikme_Aralık
187	RL_gecikme_StandartSapma	217	SS_gecikme_Entropi
188	RL_gecikme_Aralık	218	SS_gecikme_Çarpıklık
189	RL_gecikme_Entropi	219	SS_gecikme_Basıklık
190	RL_gecikme_Çarpıklık	220	SS_gecikme_KolmogorovKarmaşıklığı
191	RL_gecikme_Basıklık	221	SS_gecikme_LempelZivKarmaşıklığı
192	RL_gecikme_KolmogorovKarmaşıklığı	222	SS_gecikme_ÖrnekEntropi
193	RL_gecikme_LempelZivKarmaşıklığı	223	SS_gecikme_HFBA
194	RL_gecikme_ÖrnekEntropi	224	SS_gecikme_KFBA
195	RL_gecikme_HFBA	225	SL_gecikme_Minimum
196	RL_gecikme_KFBA	226	SL_gecikme_Maksimum
197	RS_gecikme_Minimum	227	SL_gecikme_Ortalama
198	RS_gecikme_Maksimum	228	SL_gecikme_Medyan

EK 1. Klavye tuş vuruşlarından çıkarılan öznitelikler (devam).

229	SL_gecikme_StandartSapma	259	LL_uçuş_Entropi
230	SL_gecikme_Aralık	260	LL_uçuş_Çarpıklık
231	SL_gecikme_Entropi	261	LL_uçuş_Basıklık
232	SL_gecikme_Çarpıklık	262	LL_uçuş_KolmogorovKarmaşıklığı
233	SL_gecikme_Basıklık	263	LL_uçuş_LempelZivKarmaşıklığı
234	SL_gecikme_KolmogorovKarmaşıklığı	264	LL_uçuş_ÖrnekEntropi
235	SL_gecikme_LempelZivKarmaşıklığı	265	LL_uçuş_HFBA
236	SL_gecikme_ÖrnekEntropi	266	LL_uçuş_KFBA
237	SL_gecikme_HFBA	267	LR_uçuş_Minimum
238	SL_gecikme_KFBA	268	LR_uçuş_Maksimum
239	SR_gecikme_Minimum	269	LR_uçuş_Ortalama
240	SR_gecikme_Maksimum	270	LR_uçuş_Medyan
241	SR_gecikme_Ortalama	271	LR_uçuş_StandartSapma
242	SR_gecikme_Medyan	272	LR_uçuş_Aralık
243	SR_gecikme_StandartSapma	273	LR_uçuş_Entropi
244	SR_gecikme_Aralık	274	LR_uçuş_Çarpıklık
245	SR_gecikme_Entropi	275	LR_uçuş_Basıklık
246	SR_gecikme_Çarpıklık	276	LR_uçuş_KolmogorovKarmaşıklığı
247	SR_gecikme_Basıklık	277	LR_uçuş_LempelZivKarmaşıklığı
248	SR_gecikme_KolmogorovKarmaşıklığı	278	LR_uçuş_ÖrnekEntropi
249	SR_gecikme_LempelZivKarmaşıklığı	279	LR_uçuş_HFBA
250	SR_gecikme_ÖrnekEntropi	280	LR_uçuş_KFBA
251	SR_gecikme_HFBA	281	LS_uçuş_Minimum
252	SR_gecikme_KFBA	282	LS_uçuş_Maksimum
253	LL_uçuş_Minimum	283	LS_uçuş_Ortalama
254	LL_uçuş_Maksimum	284	LS_uçuş_Medyan
255	LL_uçuş_Ortalama	285	LS_uçuş_StandartSapma
256	LL_uçuş_Medyan	286	LS_uçuş_Aralık
257	LL_uçuş_StandartSapma	287	LS_uçuş_Entropi
258	LL_uçuş_Aralık	288	LS_uçuş_Çarpıklık

EK 1. Klavye tuş vuruşlarından çıkarılan öznitelikler (devam).

289	LS_uçuş_Basıklık	319	RL_uçuş_LempelZivKarmaşıklığı
290	LS_uçuş_KolmogorovKarmaşıklığı	320	RL_uçuş_ÖrnekEntropi
291	LS_uçuş_LempelZivKarmaşıklığı	321	RL_uçuş_HFBA
292	LS_uçuş_ÖrnekEntropi	322	RL_uçuş_KFBA
293	LS_uçuş_HFBA	323	RS_uçuş_Minimum
294	LS_uçuş_KFBA	324	RS_uçuş_Maksimum
295	RR_uçuş_Minimum	325	RS_uçuş_Ortalama
296	RR_uçuş_Maksimum	326	RS_uçuş_Medyan
297	RR_uçuş_Ortalama	327	RS_uçuş_StandartSapma
298	RR_uçuş_Medyan	328	RS_uçuş_Aralık
299	RR_uçuş_StandartSapma	329	RS_uçuş_Entropi
300	RR_uçuş_Aralık	330	RS_uçuş_Çarpıklık
301	RR_uçuş_Entropi	331	RS_uçuş_Basıklık
302	RR_uçuş_Çarpıklık	332	RS_uçuş_KolmogorovKarmaşıklığı
303	RR_uçuş_Basıklık	333	RS_uçuş_LempelZivKarmaşıklığı
304	RR_uçuş_KolmogorovKarmaşıklığı	334	RS_uçuş_ÖrnekEntropi
305	RR_uçuş_LempelZivKarmaşıklığı	335	RS_uçuş_HFBA
306	RR_uçuş_ÖrnekEntropi	336	RS_uçuş_KFBA
307	RR_uçuş_HFBA	337	SS_uçuş_Minimum
308	RR_uçuş_KFBA	338	SS_uçuş_Maksimum
309	RL_uçuş_Minimum	339	SS_uçuş_Ortalama
310	RL_uçuş_Maksimum	340	SS_uçuş_Medyan
311	RL_uçuş_Ortalama	341	SS_uçuş_StandartSapma
312	RL_uçuş_Medyan	342	SS_uçuş_Aralık
313	RL_uçuş_StandartSapma	343	SS_uçuş_Entropi
314	RL_uçuş_Aralık	344	SS_uçuş_Çarpıklık
315	RL_uçuş_Entropi	345	SS_uçuş_Basıklık
316	RL_uçuş_Çarpıklık	346	SS_uçuş_KolmogorovKarmaşıklığı
317	RL_uçuş_Basıklık	347	SS_uçuş_LempelZivKarmaşıklığı
318	RL_uçuş_KolmogorovKarmaşıklığı	348	SS_uçuş_ÖrnekEntropi

EK 1. Klavye tuş vuruşlarından çıkarılan öznitelikler (devam).

349	SS_uçuş_HFBA
350	SS_uçuş_KFBA
351	SL_uçuş_Minimum
352	SL_uçuş_Maksimum
353	SL_uçuş_Ortalama
354	SL_uçuş_Medyan
355	SL_uçuş_StandartSapma
356	SL_uçuş_Aralık
357	SL_uçuş_Entropi
358	SL_uçuş_Çarpıklık
359	SL_uçuş_Basıklık
360	SL_uçuş_KolmogorovKarmaşıklığı
361	SL_uçuş_LempelZivKarmaşıklığı
362	SL_uçuş_ÖrnekEntropi
363	SL_uçuş_HFBA
364	SL_uçuş_KFBA
365	SR_uçuş_Minimum
366	SR_uçuş_Maksimum
367	SR_uçuş_Ortalama
368	SR_uçuş_Medyan
369	SR_uçuş_StandartSapma
370	SR_uçuş_Aralık
371	SR_uçuş_Entropi
372	SR_uçuş_Çarpıklık
373	SR_uçuş_Basıklık
374	SR_uçuş_KolmogorovKarmaşıklığı
375	SR_uçuş_LempelZivKarmaşıklığı
376	SR_uçuş_ÖrnekEntropi
377	SR_uçuş_HFBA
378	SR_uçuş_KFBA

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER	
Adı – Soyadı:	**Bahar DEMİR**

EĞİTİM BİLGİLERİ	
2018 -	**Trakya Üniversitesi (Yüksek Lisans)** Hesaplamalı Bilimler
2010 – 2017	**Tekirdağ Namık Kemal Üniversitesi (Hazırlık + Lisans)** Matematik
2006 – 2010	**Kırklareli Atatürk Anadolu Lisesi** Fen Bilimleri