

**BAŐKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**PARKİNSON HASTALIĐININ TUŐ VURUŐ DATASI  
KULLANILARAK TESPİTİ**

**NEZİF TAMSON**

YÜKSEK LİSANS TEZİ

2018

**PARKİNSON HASTALIĐININ TUŐ VURUŐ DATASI  
KULLANILARAK TESPİTİ**

**DETECTION OF PARKINSON DISEASE BY USING  
KEYSTROKE DATA**

**NEZİF TAMSON**

Başkent Üniversitesi  
Lisansüstü Eğitim Öğretim ve Sınav YönetmeliĐinin  
ELEKTRİK-ELEKTRONİK MühendisliĐi Anabilim Dalı İin ÖngördüĐü  
YÜKSEK LİSANS TEZİ  
olarak hazırlanmıŐtır.  
2018

“Parkinson Hastalığının Tuş Vuruş Datası Kullanılarak Tespiti” isimli bu çalışma jürimiz tarafından, 26/12/2018 tarihinde, **ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI** 'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Başkan : Doç. Dr. Hasan Şakir BİLGE

Üye (Danışman) : Dr. Öğr. Üyesi Derya YILMAZ

Üye : Dr. Öğr. Üyesi Selda GÜNEY

ONAY

11/01/2019

Prof. Dr. Faruk ELALDI  
Fen Bilimleri Enstitüsü Müdürü



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

Tarih:11/01/2019

Öğrencinin Adı, Soyadı : Nezif TAMSON

Öğrencinin Numarası : 21620212

Anabilim Dalı : Elektrik-Elektronik Mühendisliği

Programı : Yüksek Lisans Programı

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Derya YILMAZ

Tez Başlığı : Parkinson hastalığının tuş vuruş datası kullanılarak tespiti

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 81 sayfalık kısmına ilişkin, 11/01/2019 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 13'dür.

Uygulanan filtrelemeler:

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

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

Öğrenci İmzası:.....

Onay

11/01/2019

Dr. Öğr. Üyesi Derya YILMAZ

## **TEŐEKKÜR**

Tez danıőmanım Sayın Dr. Derya Yılmaz'a, bana gvenerek benimle tez alıőması yapmayı kabul ettiėi iin,

Eėitimim sresince bana deėerli bilgilerini aktaran tm Baőkent niversitesi Elektrik – Elektronik Mhendisliėi oėretim elemanlarına,

Ve hayatım boyunca bana gvenip, maddi ve manevi Őekilde her zaman yanımda olan aileme

Sonsuz teőekkr ederim.

## ÖZ

### PARKİNSON HASTALIĞININ TUŞ VURUŞ DATASI KULLANILARAK TESPİTİ

Nezif Tamson

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Parkinson; genellikle ellerde ve ayaklarda titreme, kaslarda sertlik, hareketlerde zorluk ve duruş bozukluğu belirtileriyle ortaya çıkan ve kademeli olarak ilerleyen bir sinir sistemi hastalığıdır. Parkinsonun teşhisi için yapılan çalışmalarda kuvvet ve ivme sensörleri, salınım fazları ve eylemsizlik ölçümleri kullanılarak; yürüme, duruş ve hareket bozukluğu karakteristikleri incelenmiştir. Son yıllarda mevcut teşhis yöntemlerine göre üstün yönleri nedeniyle bu hastalığın tespitinde, insan-bilgisayar etkileşiminden hareketle; Parkinson hastalarının klavye kullanımına ilişkin tuş vuruş dinamiklerinden elde edilen verilerin değerlendirilmesi önem kazanmıştır. Bu çalışmada, sağlıklı ve Parkinson hastası bireylerden günlük bilgisayar kullanımı sırasında kaydedilen tuş vuruş verileri analiz edilerek, Parkinson hastalığının tespiti üzerinde çalışılmıştır. Verilerden; yüksek dereceli momentler, entropiler, simetri bozukluğu ve istatistiksel nicelikler başta olmak üzere toplam 14 özellik hesaplanmıştır. Ayırt edici özellikler, istatistiksel testler ve Rastgele Orman (RO) algoritması kullanılarak, Destek Vektör Makinesi (DVM) ve k En Yakın Komşu (kEYK) sınıflayıcılarına uygulanmıştır. Sınıflayıcılar; eğitim ve test oranlarının 50-50 ve 30-70 olduğu iki farklı durumda, 646 ve 515 veri içeren iki veri kümesi için çalıştırılmıştır. En yüksek doğruluk oranları kEYK sınıflayıcısıyla; 646 veri için 83,78% eğitim ve 80,15% test; 515 veri için de 86,64% eğitim ve 82,4% test olarak elde edilmiştir. Çalışmanın sonuçları, tuş vuruş dinamiklerinin Parkinson hastalığının tespitinde kullanılabileceğini göstermiştir.

**ANAHTAR SÖZCÜKLER:** Parkinson hastalığı; tuş vuruş dinamikleri, insan-bilgisayar etkileşimi, özellik çıkarma, sınıflama.

**Danışman:** Dr. Öğr. Üyesi Derya YILMAZ, Başkent Üniversitesi, Elektrik – Elektronik Mühendisliği Anabilim Dalı

## **ABSTRACT**

### **DETECTION OF PARKINSON DISEASE BY USING KEYSTROKE DATA**

Nezif Tamson

Baskent University Institute of Science and Engineering

Department of Electrical and Electronics Engineering

Parkinson's Disease (PD) is a neurological movement disorder that occurs in the hands and feet with tremor, rigidity, slowing of movements and difficulty walking and postural instability. Generally; the measures from force sensors, accelerometers and inertia measurement units used to gain informations about gait, posture and disorderly movements have been studied for analyzing the PD's characteristics. In the last few years, due to the superior aspects of the current diagnostic methods, based on human-computer interaction; evaluation of data obtained from keystroke dynamics of keyboard use of Parkinson's patients has gained importance. In this study total 14 features, including asymmetry, entropies, high degree momentums and statistical quantities were calculated from datas and have been studied to determine the PD. All these significant features, statistical tests and Random Forest algorithms were used and applied to the inputs of two-class Support Vector Machines (SVM) and k Nearest Neighbor (kNN). The classifier accuracies were both found for training and testing in terms of 50-50 & 30-70 respectively. These results are listed for both 646 and 515 records. The obtained features were evaluated in four different cases. In all cases, the highest test accuracy is 80,15% (training: 83,78%) for 646 records and 82,4% (training: 86,64%) for 515 records, found by kNN classifier. These results shown that keystroke datas are able to used for PD diagnosing instead of other sensor measures.

**KEYWORDS:** Parkinson's disease, keystroke dynamics, feature selection, human – computer interaction, classification

**Advisor:** Dr. Derya YILMAZ, Baskent University, Department of Electrical and Electronics Engineering.

# İÇİNDEKİLER LİSTESİ

TEŞEKKÜR.....	
ÖZET.....	i
ABSTRACT.....	ii
İÇİNDEKİLER LİSTESİ.....	iii
ŞEKİLLER LİSTESİ.....	v
ÇİZELGELER LİSTESİ.....	ix
SİMGELER VE KISALTMALAR LİSTESİ.....	x
<b>1. GİRİŞ.....</b>	<b>1</b>
1.1 Tuş Vuruş Verileri Kullanılarak Yapılan Çalışmalar ve Tarihsel Gelişimi.....	2
1.2 Parkinson Hastalığında Kullanılan Teknikler ve Tarihsel Gelişimi.....	4
1.2.1 Uzman klinisyenler (nöroloji uzmanları).....	4
1.2.2 Birleşik Parkinson Hastalığı Derecenelendirme Ölçeği (BPHDÖ).....	4
1.2.3 İvmeölçer (Hız Ölçer ) sensörü.....	5
1.2.4 Kuvvet sensörü.....	5
1.2.5 Jiro sensörü (Jiroskop).....	6
1.2.6 Parmakla dokunma testi ve bilgisayar klavyesi etkileşimi.....	6
1.3 Çalışmanın Genel Yapısı.....	6
<b>2. MATERYAL VE METOTLAR.....</b>	<b>7</b>
2.1. Veri Seti.....	7
2.2. Çıkarılan Özellikler.....	13
2.2.1 Asimetri hesabı (Simetri Bozukluğu).....	13
2.2.2 Histogram alanları hesaplanması.....	14
2.2.3 Ortalama, Medyan, Standart Sapma, Varyans, Çarpıklık ve Basıklık parametreleri.....	16
2.2.4 Durum uzayı ( $X_n$ vs $X_{n+1}$ ).....	18
2.2.5 İlk fark (first differencing) alınarak elde edilen durum uzayı.....	19
2.2.6 Entropi.....	20
2.2.7 Ortalama Ortak Bilgi (Average Mutual Information).....	23
2.2.8 Yerleştirme Boyutu (Embedding Dimension).....	25



2.3. Özellik Seçimi.....	26
2.3.1 Varyans Analizi (ANOVA).....	26
2.3.2 Rastgele Orman (RO) (Random Forest- RF) Algoritması.....	27
2.4. Sınıflama Algoritmaları.....	28
2.4.1 Destek Vektör Makinesi (DVM)( Support Vector Machine (SVM's)).....	28
2.4.2 K En Yakın Komşu(kEYK) ( k Nearest Neighbor (k-NN) ).....	35
<b>3. SONUÇLAR.....</b>	<b>37</b>
3.1. İstatistiksel Analiz Sonuçları.....	37
3.2. Sınıflayıcı Sonuçları.....	60
3.2.1 Destek Vektör Makinesi sınıflayıcısı sonuçları.....	65
3.2.2 K En Yakın Komşu sınıflayıcısı sonuçları.....	71
<b>4. SONUÇ.....</b>	<b>75</b>
<b>5. KAYNAKLAR LİSTESİ.....</b>	<b>78</b>

## ŞEKİLLER LİSTESİ

Şekil 1.1	Çalışmanın genel yapısı.....	7
Şekil 2.1	Bir katılımcıya ait detaylı kişisel bilgilerin olduğu dosyanın ekran görüntüsü.....	9
Şekil 2.2	Bir katılımcıya ait zaman ve yön verileri.....	10
Şekil 2.3	Basılı tutma, gecikme ve geçiş sürelerini gösteren diyagram .....	11
Şekil 2.4	Parkinson hastası bir katılımcının basılı tutma verisi.....	12
Şekil 2.5	Sağlıklı bir katılımcının basılı tutma verisi .....	13
Şekil 2.6	Parkinson hastası bir katılımcının basılı tutma sürelerinin Histogram'ı16	
Şekil 2.7	Sağlıklı bir katılımcının basılı tutma sürelerinin Histogram'ı.....	16
Şekil 2.8	Parkinson hastası bir katılımcıdan elde edilen durum uzayı.....	18
Şekil 2.9	Sağlıklı bir katılımcıdan elde edilen durum uzayı.....	18
Şekil 2.10	Parkinson hastası bir katılımcıdan elde edilen ilk fark durum uzayı....	19
Şekil 2.11	Sağlıklı bir katılımcıdan elde edilen ilk fark durum uzayı.....	21
Şekil 2.12	Bir katılımcıya ait I(T) fonksiyonu grafiği.....	24
Şekil 2.13	Destek Vektör Makinesi Yapısı.....	29
Şekil 2.14	2 Sınıflı Veri Setini Ayıran Farklı Düzlemlere İlişkin Örnek.....	30
Şekil 2.15	Destek Vektör Makinesi için 2 sınıflı Problem Örneği.....	30
Şekil 2.16	2 Sınıflı Problem İçin Doğrusal Ayrılama Durumu.....	33
Şekil 2.17	Doğrusal Olmayan Haritalama Yaklaşımı Örneği .....	34
Şekil 3.1	646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan LL-RR Asimetri, LR-RL Asimetri ve Shannon Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.....	39
Şekil 3.2	646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Ortalama, Medyan ve Standart Sapma özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	39
Şekil 3.3	646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Varyans özelliğinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	40
Şekil 3.4	646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Yaklaşık Entropi ve İlk Fark IV. bölge nokta sayılarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.....	40
Şekil 3.5	646 kayda göre basılı tutma sürelerinin Histogram alanlarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.....	42
Şekil 3.6	515 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Varyans özelliğinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	43

Şekil 3.7	515 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Ortalama, Medyan ve Standart Sapma özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	43
Şekil 3.8	515 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan LL-RR Asimetri, LR-RL Asimetri, Shannon Entropi ve İlk Fark IV. Bölge özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.....	43
Şekil 3.9	515 kayda göre basılı tutma sürelerinin Histogram alanları.....	44
Şekil 3.10	646 kayda göre gecikme sürelerinde ayırt edici olarak bulunan durum uzayı 1., 2., 3. ve 4. bölge nokta sayılarının Parkinson hastası ve sağlıklı bireyler için ortalama değerleri.....	46
Şekil 3.11	646 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Basıklık ve Ortak Bilgi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	46
Şekil 3.12	646 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Çarpıklık ve Yaklaşık Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.....	46
Şekil 3.13	646 kayda göre gecikme sürelerinin Histogram alanları.....	48
Şekil 3.14	515 kayda göre gecikme sürelerinde ayırt edici olarak bulunan 1., 2., 3. ve 4. bölge nokta sayılarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	50
Şekil 3.15	515 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Ortak Bilgi, Örnek Entropi ve Basıklık özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	50
Şekil 3.16	515 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Yaklaşık Entropi ve Çarpıklık özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	50
Şekil 3.17	515 kayda göre gecikme sürelerinin Histogram alanları.....	51
Şekil 3.18	646 kayda göre geçiş sürelerinde ayırt edici olarak bulunan 1., 2., 3. ve 4. bölge nokta sayılarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.....	53
Şekil 3.19	646 kayda göre geçiş sürelerinde ayırt edici olarak bulunan Çarpıklık, Basıklık ve Yaklaşık Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	53
Şekil 3.20	646 kayda göre geçiş sürelerinin Histogram alanları.....	55
Şekil 3.21	515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan LL-RR Asimetri özelliğinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.....	57
Şekil 3.22	515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan İlk Fark durum uzayı IV ve III.bölge nokta sayıları ile durum uzayı 1., 2., 3. ve 4.bölge nokta sayılarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	57

Şekil 3.23	515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan Çarpıklık ve Basıklık özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	58
Şekil 3.24	515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan Örnek Entropi ve Yaklaşık Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri .....	58
Şekil 3.25	515 kayda göre geçiş sürelerinin Histogram alanları.....	60
Şekil 3.26	646 kayda göre $P \leq 0,001$ çıkan 13 özelliğin önem derecelerini göstermektedir.....	64
Şekil 3.27	515 kayda göre $P \leq 0,001$ çıkan 14 özelliğin önem derecelerini göstermektedir.....	64
Şekil 3.28	Rasgele Orman algoritması ile 646 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.....	71
Şekil 3.29	Rasgele Orman algoritması ile 515 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.....	73

## ÇİZELGELER LİSTESİ

Çizelge 2.1	Katılımcıların detaylı kişisel bilgilerinin olduğu klasörün veri elemanları ve detayları.....	8
Çizelge 2.2	Bir katılımcının bilgisayar klavyesi ile olan etkileşiminden elde edilen dinamikler.....	9
Çizelge 3.1	Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre basılı tutma sürelerinin ANOVA sonuçları.....	38
Çizelge 3.2	Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre basılı tutma sürelerinin Histogram alanlarının ANOVA sonuçları.....	41
Çizelge 3.3	Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre basılı tutma sürelerinin ANOVA sonuçları.....	42
Çizelge 3.4	Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre basılı tutma sürelerinin Histogram alanlarının ANOVA sonuçları.....	44
Çizelge 3.5	Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre gecikme sürelerinin ANOVA sonuçları.....	45
Çizelge 3.6	Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre gecikme sürelerinin Histogram alanlarının ANOVA sonuçları.....	47
Çizelge 3.7	Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre gecikme sürelerinin ANOVA sonuçları.....	49
Çizelge 3.8	Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre gecikme sürelerinin Histogram alanlarının ANOVA sonuçları.....	51
Çizelge 3.9	Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre geçiş sürelerinin ANOVA sonuçları.....	52
Çizelge 3.10	Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre geçiş sürelerinin Histogram alanlarının ANOVA sonuçları.....	54
Çizelge 3.11	Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre geçiş sürelerinin ANOVA sonuçları.....	56
Çizelge 3.12	Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre geçiş sürelerinin Histogram alanlarının ANOVA sonuçları.....	59
Çizelge 3.13	646 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde Destek Vektör Makinesi sınıflayıcısına verilerek elde edilen sonuçları.....	61
Çizelge 3.14	515 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde Destek Vektör Makinesi sınıflayıcısına verilerek elde edilen sonuçları.....	62
Çizelge 3.15	646 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde Destek Vektör Makinesi sınıflayıcısına verilerek elde edilen sonuçları.....	62
Çizelge 3.16	515 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde Destek Vektör Makinesi sınıflayıcısına verilerek elde edilen sonuçları.....	63

Çizelge 3.17	646 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte Destek Vektör Makinesi sınıflayıcısına verilerek elde edilen sonuçları.....	63
Çizelge 3.18	515 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte Destek Vektör Makinesi sınıflayıcısına verilerek elde edilen sonuçları.....	63
Çizelge 3.19	Rasgele Orman algoritması ile 646 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen Destek Vektör Makinesi sınıflayıcısı doğruluk oranları.....	65
Çizelge 3.20	Rasgele Orman algoritması ile 515 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen Destek Vektör Makinesi sınıflayıcısı doğruluk oranları.....	66
Çizelge 3.21	646 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde k En Yakın Komşu sınıflayıcısına verilerek elde edilen sonuçları.....	67
Çizelge 3.22	515 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde k En Yakın Komşu sınıflayıcısına verilerek elde edilen sonuçları.....	68
Çizelge 3.23	646 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde k En Yakın Komşu sınıflayıcısına verilerek elde edilen sonuçları.....	68
Çizelge 3.24	515 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde k En Yakın Komşu sınıflayıcısına verilerek elde edilen sonuçları.....	69
Çizelge 3.25	646 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte k En Yakın Komşu sınıflayıcısına verilerek elde edilen sonuçları.....	69
Çizelge 3.26	515 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte k En Yakın Komşu sınıflayıcısına verilerek elde edilen sonuçları.....	69
Çizelge 3.27	Rasgele Orman algoritması ile 646 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.....	69
Çizelge3.28	Rasgele Orman algoritması ile 515 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.....	76
Çizelge 3.29	646 kayıt için elde edilen en yüksek test doğruluk oranının (80,15 %) konfüzyon matrisi.....	78
Çizelge 3.30	515 kayıt için elde edilen en yüksek test doğruluk oranının (80,15 %) konfüzyon matrisi.....	78

## SİMGELER VE KISALTMALAR LİSTESİ

Abs	Mutlak Değer (Abstract)
LL	Soldan Sola (Left to Left)
LR	Soldan Sağa (Left to Right)
RL	Sağdan Sola (Right to Left)
RR	Sağdan Sağa (Right to Right)
GRF	Ground Reaction Force (Yer Reaksiyon Kuvveti)
ApEn	Approximate Entropy (Yaklaşık Entropi)
SampEn	Sample Entropy (Örnek Entropi)
S	Boşluk Tuşu (Space)
HT	Hold Time (Basılı Tutma Süresi)
LT	Latency Time (Gecikme Süresi)
FT	Flight Time (Geçiş Süresi)
FNN	False Nearest Neighbor (En Yakın Yanlış Komşu)
SVM	Support Vector Machine (Destek Vektör Makinesi)
kNN	k Nearest Neighbor (k En Yakın Komşu)
BPHDÖ	Birleşik Parkinson Hastalığı Derecelendirme Ölçeği & Unified Parkinson's Disease Rating Scale (UPDRS)
kEYK	k En Yakın Komşu
DVM	Destek Vektör Makinesi
EYYK	En Yakın Yanlış Komşu
RO	Rastgele Orman Algoritması

## 1. GİRİŞ

Teknoloji çağı olarak adlandırılan içinde yaşadığımız bu çağda özellikle bilgisayar kullanımı gittikçe artmaktadır. Hayatımızın her alanına giren bilgisayarlar ile bu kullanımlardan elde edilecek bilgiler her geçen gün önem kazanmaktadır. Tuş vuruş dinamikleri bilgisayarı kullanan kişinin klavyede yazarken alışkanlık haline getirdiği kalıpları ve yazma ritimlerini içerir. Bu ritim ve yazma desenleri tıpkı el yazısı ve imzada olduğu gibi kişiye has bir yapısı vardır. Tuş vuruş dinamikleri bilgisayarı kullanan kişinin ne yazdığına değil nasıl yazdığına ve ne kadar sürede yazdığına odaklanır. Literatür çalışmalarına bakıldığında bugüne kadar bu kalıplardan ve ritimlerden yararlanılarak kimlik doğrulama gerektiren sektörler ve sistemlerde [1], duygu durumu belirlenmesinde [2], stres seviyesinin belirlenmesinde [3], yaş tespitinde [4] ve bazı hastalıkların tespitinde [5] kullanılmıştır. Literatür incelendiğinde tuş vuruş dinamiklerinin analizi sabit/yapılandırılmış metin veya dinamik/serbest metin olarak ikiye ayrılır. Önceden tanımlanmış aynı metnin katılımcı tarafından sisteme girilmesi sabit metin tuş vuruş dinamiği olarak tanımlanmaktadır. Katılımcının herhangi bir komuta bağlı kalmadan kendilerinin yazmış oldukları metinler ise bağımsız/serbest metin olarak tanımlanmaktadır.

Parkinson hastalığı ise kademeli olarak ilerleyen sinir sistemi hastalığıdır. Sinir sistemi hastalıkları içerisinde Alzheimer' dan sonra en yaygın görülen ikinci hastalıktır. Bu hastalık artan yaş ile birlikte etkisini daha da çok hissettirmektedir. Genellikle 65 yaş üstü kişileri etkilemektedir ancak bu yaşın altında görülme ihtimali de mevcuttur. Amerika Birleşik Devletlerinde yaklaşık bir milyon, dünyada ise yaklaşık 10 milyon kişinin parkinson hastası olduğu bilinmektedir. Parkinson hastalığı beyindeki dopamine (DA) seviyesinin azalmasıyla ortaya çıkmaktadır. Bu hastalık yavaşça ilerlemektedir. Genellikle ellerde ve ayaklarda titreme, duruş bozukluğu, yürüme ve denge problemleri, kaslarda katılık şeklinde ortaya çıkmaktadır. Bunun haricinde birçok motor ve motor olmayan semptomlar da mevcuttur. Bu semptomların ilerlemesiyle birlikte hastalığın günlük yaşama etkisi artmaktadır ve kişinin yaşam kalitesini düşürmektedir. Bugüne kadar yapılan çalışmalarda araştırmacılar kuvvet sensörü, ivmeölçer, jiroskop gibi ölçüm araçlarını kullanarak parkinson hastalığını tespit etmeye çalıştılar. Parkinson hastalığının günümüzde tespiti için kullanılan yöntemler göz önüne alındığında kliniksel muayeneler ve sensörler aracılığı ile



yapılan analizler hem maddi hem de manevi şekilde hastayı zorlamaktadır. Mevcut tıbbi uygulamalar Parkinson hastalığındaki motor semptomların işaretlerini anlamakta yetersiz kalmaktadır ve bu hastalığın düzenli olarak izlenmesi için yeterli araca ve yazılıma sahip değildirler. Diğer bir yaygın metot ise Birleşik Parkinson Hastalığı Derecelendirme Ölçeği'nin (UPDRS) kullanılmasıdır. Klinisyenler veya fizikçiler tarafından yapılan testlerde katılımcıların vücutlarına sensörler takılması ve bazı komutların (TUG Test gibi) uygulanması gerektiğinin katılımcıya söylenmesi üzerine katılımcının heyecanlanmasına ve elde edilecek ölçümlerin etkilenmesine sebep olmaktadır. Özellikle Parkinson hastası olmayıp da bundan şüphe duyan bireyler bu testlerden daha çok etkilenmektedirler.

Parkinson hastalığının tespitinde kullanılan yukarıda bahsedilen dezavantajlar göz önüne alındığında ve insan bilgisayar etkileşiminin günümüzdeki kullanımının artmasıyla birlikte, bu alan için de yeni bir yaklaşım ortaya çıkmıştır. Yapılan bu çalışmada serbest metinler kullanan katılımcıların tuş vuruş dinamikleri incelenmiş ve bu kişilerde Parkinson hastalığının olup olmadığı anlaşılmaya çalışılmıştır. Bu çalışmanın avantajı katılımcıların herhangi bir özel ekipman giymesine veya tıbbi bir gözetime maruz kalmasına gerek duyulmaması ve uygulayıcının deneyimine ve becerisine bağlı olmamasıdır. İnsan bilgisayar klavyesi etkileşimi, Parkinson hastalığının tespit edilmesinin haricinde Karpal Tünel Sendromu (KTS) ve göz takip arayüzü ile ALS hastalığının tespitinde de kullanılmaktadır [15,16]. İnsan bilgisayar klavyesi etkileşiminden yararlanılarak Parkinson hastalığının tespiti daha önce iki farklı ekip tarafından çalışılmıştır [17, 18].

### **1.1 Tuş Vuruş Verileri Kullanılarak Yapılan Çalışmalar ve Tarihsel Gelişimi**

Geçmiş 2. Dünya Savaşına kadar dayanan tuş vuruş dinamikleri tabanlı biyometrik kimlik doğrulama sistemleri, klavyede yazma karakteristiğini ölçen, taklit etmenin zor ve yapısı itibari ile eşsiz olduğu düşünülen sistemlerdir.

Zeki Özen ve arkadaşları; katılımcılara klavyeden girmeleri için verilen şifreyi yazarken arka arkaya üç kere basılan CAPS LOCK tuşunu incelemişler ve model olarak ise Yapay Sinir Ağlarının Geri Yayılım Algoritması, Radyal Tabanlı Fonksiyon ve Öz İlişkili Algoritmasını kullanmışlardır [1]. Ayrıca 5 kat çapraz doğrulama da uygulanmıştır. Sonuç olarak Geri Yayılım Algoritması ile 94.40%, Radyal Tabanlı

Fonksiyon ile 91.01% ve Öz İlişkili Algoritma ile 94.69% doğruluk oranı elde etmişlerdir.

Avar Pentel; 2017 yılında yapmış olduğu çalışmada; bilgisayar klavyesi ve faresini kullanarak katılımcıların duygu durumlarını anlamaya çalışmıştır [2]. Duygu durumlarının, izlenen filmler karşısında değiştiğini düşünerek katılımcılara 4 farklı konuda film kesitleri izletilerek sonrasında katılımcılardan izlenen film kesitleri sonrası katılımcılardan milli marşın 3 farklı versiyonunun bilgisayar klavyesi ile yazılması istenmiştir. Basılı tutma süresi ve geçiş süresi dinamiklerinin ortalamaları, standart sapmaları ve silme tuşunun kaç kere tuşlandığı değerlendirilerek duygu durumu ayırt edilmeye çalışılmıştır. Lojistik Regresyon, Destek Vektör Makineleri (DVM), k En Yakın Komşu (KEYK (k Nearest Neighbor - kNN)), C4.5 ve Rastgele Orman (RO (Random Forest)) sınıflandırıcılarını kullanarak sırası ile 71%, 67%, 84%, 77% ve 80% doğruluk oranlarını elde etmiştir.

Avar Pentel ise 2018 yılında yapmış olduğu çalışmada , 6 farklı yaş grubundan toplam 7119 kişiden tuşlara vurma verileri toplamıştır. Ortalama basılı tutma süreleri, ortalama geçiş süreleri, kelimeler arasındaki ortalama durma sürelerine odaklanmış ve makine dili öğrenme algoritmalarından DVM, RO ve Lojistik Regresyon gibi sınıflayıcılarını kullanarak en düşük 82% ve en yüksek 92% doğruluk oranlarını elde etmiştir [4].

Shing Hon Lau, 2018 yılında yapmış olduğu çalışmada; stresin insan davranışlarını derinden etkilediğini düşünmüş ve stresli olunan durumlarda hataların daha çok yapıldığını farketmiştir [3]. Yapmış oldukları çalışmada katılımcılardan ilk olarak dinlenir durumda iken klavyeden bir takım yazılar yazmalarını istemiş ve bunlar kaydedilmiş, ardından katılımcılara stres yaptırarak bir takım kayıtlar daha almışlar ve son olarak tekrar dinlenme durumunda iken katılımcılara klavyeden bir takım yazılar yazdırmış ve bunları kayıt altına almıştır. 3 durumda da yazılan yazı daha önceden belirlenmiş ve kullanıcılara söylenmiştir. Tuş vuruş dinamiklerinin ortalama ve medyan değerleri hesaplanmış ve çıkarılan özellikleri 3 farklı sınıflayıcı ile test etti. Kullandığı sınıflayıcılar RO, DVM ve Lojistik Regresyon'dur. En yüksek doğruluk oranını 93.67% ile RO'ya aittir. En düşük doğruluk oranı ise 67.76% ile Lojistik Regresyon' a aittir.

## **1.2 Parkinson Hastalığının Tespitinde Kullanılan Teknikler ve Tarihsel Gelişimi**

Nöroloji uzmanları genellikle kliniklerinde yaptıkları testlerle veya muayene edilen kişinin EEG ve EMG'lerini inceleyerek Parkinson hastası olup olmadığına karar vermeye çalıştılar [5].

### **1.2.1 Birleşik Parkinson Hastalığı Derecelendirme Ölçeği (BPHDÖ)**

Birleşik Parkinson Hastalığı Derecelendirme Ölçeği (BPHDÖ) ilk olarak 1980'lerde ortaya çıkmış ve kliniksel olarak parkinson hastalığının değerlendirilmesinde en yaygın olarak kullanılan teknik olmuştur [6]. Hemen hemen her durumda Parkinson hastalığının tedavisinin değerlendirilmesi için yapılan testlerde BPHDÖ tedaviye bağlı yararları belirlemek için kullanılan bir ölçektir. BPHDÖ'nin amacı yaşın, ilaç tedavisinin veya hastalığın şiddeti ne olursa olsun kapsamlı ve pratik bir şekilde kullanılabilen bir ölçek olmasıdır. Bu ölçek 0-5 arasında olup hastalık ona göre derecelendirilmektedir.

20 yılın ardından 2001 yılında MDS tarafından içeriği değiştirilmiş, yeni sorular eklenmiş ve yeni versiyonu 2008 yılında yayınlanmıştır. Klinisyenler tarafından sorulan sorular sorulup katılımcılar tarafından cevaplanmaktadır. Yaklaşık 30 dakika süren soru cevap şeklinde bir testtir [7].

### **1.2.2 İvmeölçer (Hız Ölçer) sensörü**

Parkinson hastalığının tespiti için kullanılan diğer bir yöntem ise katılımcıların vücutlarına 3 eksenli ivmeölçer sensörü takılmasıdır. Kişilerin vücutlarının farklı bölgelerine yerleştirilen sensörlerden x,y ve z eksenlerinden elde edilen ivme ölçümleri değerlendirilerek hastalık tespit edilmeye çalışılmıştır. Parkinson hastalığının bilinen semptomlarından bir tanesi "Freezing of Gait" yani yürüme esnasında gerçekleşen donmalardır. Özellikle parkinson hastalarında dar bir yerde yürürken veya köşeleri dönerken bu donmalar daha sık meydana gelmektedir. Bu gibi durumlarda parkinson hastası bir bireyin hızı sağlıklı bir kişiye göre daha az olacaktır.

Evanthia E. Tripoliti ve arkadaşlarının 2014 yılında yapmış oldukları çalışmada 6 tane ivmeölçer ve 2 tane jiroskop kullanılmış ve 96.11% doğruluk oranı elde edilmiştir [8]. Çalışma sırasında ivmeölçerler el ve ayak bileklerine, göğüze ve bele takılmıştır. Jiroskoplar ise katılımcıların göğsüne ve beline takılmıştır.

Sinziana Mazilu ve arkadaşları ise 2015 yılında yapmış oldukları çalışmada katılımcıların el bilekleri, ayak bilekleri ve dizlerine toplamda 6 adet ivmeölçer sensörü yerleştirmişler; el ve ayak hareketlerindeki hızlardan parkinson hastalığını tespit etmeye çalışmışlardır. Bu çalışmanın sonucunda 90% doğruluk oranı elde etmişlerdir [9].

Claas Ahlrichs ve arkadaşları, 2016 yılında yapmış oldukları çalışmada; 3 eksenli sensörü katılımcıların beline takmışlar, bu sensör yardımı ile hareketlerini ve hızını inceleyerek 98.7% doğruluk oranı elde etmişlerdir [10].

### **1.2.3 Kuvvet sensörü**

Parkinson hastalığının tespitinde kullanılan yöntemlerin başında yürüyüş kinematikleri ve kinetikleri karakteristiklerinin kullanılması gelmekle birlikte ayağın altına giyilebilen ve yer reaksiyon kuvvetini (GRF: Gound Reaction Force) hesaplayabilen sensörler de kullanılmıştır. Bu sensör sayesinde GRF'nin üç eksenindeki bileşeni hesaplanmaktadır ama esas olarak dikkate alınan dik bileşendir.

Sang Hong Lee ve arkadaşları 2012 yılında yapmış oldukları çalışmada; katılımcıların ayaklarının topuk, orta ve uç kısımlarına denk gelecek şekilde toplamda 16 adet (her bir ayağa 8) üç eksenli kuvvet sensörü yerleştirmişlerdir [11]. Her bir sensöre uygulanan dik kuvvet hesaplanmış ardından her bir ayağa uygulanan toplam kuvvet hesaplanmış ve iki ayak arasındaki kuvvet farkından yola çıkarak 74.32% doğruluk oranı elde etmişlerdir.

Benoit Mariani ve arkadaşları ise 2013 yılında yapmış oldukları çalışmada katılımcının adımlama süresince topuklarının ve parmak uçlarının vuruş ve kalkışlarda maruz kaldığı kuvveti hesaplamışlardır [12]. Bu araştırmacılar yaptıkları çalışmanın doğruluk oranı hakkında bilgi vermemişlerdir.

### **1.2.4 Jiro sensörü (Jiroskop)**

Parkinson hastalığının tespitinde kullanılan diğer bir yöntem ise jiroskop sensörlerdir. Jiroskop (düzdöner) sensörleri açısal hızı ölçmektedir. Bu sensörler 3 boyutlu olup 3 adet eksenle oluşmaktadır. Bu eksenler roll (yatış açısı), pitch (yükselme alçalma adım açısı) ve yaw (sapma açısı) olmak üzere 3 boyuttan oluşur. Roll eksen, sağa ve sola yatışlardır. Pitch eksen vücudun aşağı ve yukarı doğru hareketlerini ifade

eder. Y eksenini ise sağı ve sola dönüşleri ifade eder. A. Salarian ve çalışma arkadaşları 2004 yılında yapmış oldukları araştırmalarında her bir el bileğine birer tane 3 boyutlu jiroskop yerleştirilerek bradikinezi ve titremeyi algılayarak Parkinson hastalığını tespit etmeye çalışmışlardır [13].

### **1.2.5 Parmakla dokunma testi ve bilgisayar klavyesi etkileşimi**

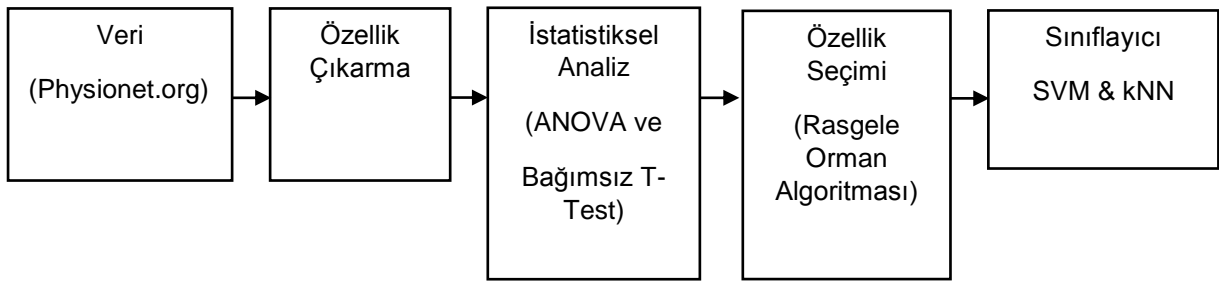
Parmakla dokunma testleri Parkinson hastalarının kas kontrolünün ve motor yeteneklerin işlevselliğinin değerlendirilmesi için kullanılabilen bir yöntemdir. Özellikle motor fonksiyonlarının parkinson hastalarının ve sağlıklı bireylerin el tercihlerinde ve el yeteneklerinde nasıl reaksiyon verdiğini görmek için tercih edilir. Parmakla dokunma testleri için çeşitli çalışmalar yapılmış ve günümüzde bu çalışmalar devam etmektedir. Araştırmacılar parmakla dokunma testi için farklı cihazlar kullanmışlardır. Bunlardan bazıları mekanik sayaçlar, elektronik anahtarlar, piyano klavyeleri, telefon ekranları ve telgraflardır [14]. Parmak hızları, tuşlara yapılan darbelerin kuvvetleri, iki tuş arasında geçen süreler gibi özellikler incelenerek hastalık tespit edilmeye çalışılmıştır. L. Giancardo ve arkadaşları [17] yapmış oldukları çalışmada sadece basılı tutma sürelerine odaklanarak 81%; Warwick R. Adams ve arkadaşları ise yapmış oldukları çalışmada basılı tutma ve gecikme sürelerine odaklanarak 98% doğruluk oranı elde etmişlerdir. Warwick R. Adams ve arkadaşlarının yapmış olduğu çalışmanın doğruluk oranının bu kadar yüksek olmasının sebebi katılımcıların hangi tuşa bastıklarını bilmeleridir. Warwick R. Adams ve arkadaşlarının yapmış olduğu çalışmada katılımcılar sırası ile 'G', 'O', 'A', 'D' tuşlarına basmışlardır [18].

### **1.2 Çalışmanın Genel Yapısı**

Yapılan bu çalışmada ise parkinson hastalığının tespiti için bilgisayar klavyesi kullanılmıştır. Bu çalışmada katılımcıların bilgisayar klavyesi ile olan etkileşimindeki zaman dinamikleri incelenerek sağlıklı bireylerin parkinson hastalarından ayırt edilmesi çalışılmıştır. Katılımcıların herhangi bir komut almadan ve herhangi bir gözlem gözetmeksizin bilgisayar klavyesi tuşlarına yapmış oldukları basmalardan elde edilen veriler kullanılmış ve Parkinson hastalığı tespit edilmeye çalışılmıştır. Bu veriler Physionet.org (Ekim 2017) veri tabanından alınmıştır. Alınan bu veriler 166 tanesi Parkinson hastası ve 50 tanesi sağlıklı olmak üzere 200' den fazla katılımcıdan elde edilmiştir Teknoloji tabanlı teşhis yöntemleri yüksek doğruluk oranı

ile sonuç vermekte, tekrarlanabilmekte ve katılımcının kullanabilmesi açısından da basit bir yapıdadırlar. Bu çalışmada da katılımcılar uygulamaya istedikleri zaman erişebilmişler, istedikleri günlük bilgisayar aktivitelerini (mail yazıp gönderme, internette bir şey arama, vb ..) herhangi bir gözleme tabi tutulmadan yapmışlardır. Kullanıcıya ait zaman verileri "Tappy" isimli yazılım uygulaması ile kaydedilmiştir. Elde edilen her bir zaman serisi için 9 farklı özellik hesaplanmıştır. Bu 9 özellik içerisinde ayırt edici olarak bulunan özellikler Gauss tabanlı makine dili öğrenme algoritması olan Destek Vektör Makinesi (Support Vector Machine (DVM) ve k En Yakın Komşu (kEYK) Sınıflayıcı girişlerine uygulanmıştır.

Şekil 1.1'de çalışmanın genel yapısı blok diyagramı ile özetlenmiştir.



Şekil 1. 1 Çalışmanın genel yapısı.

## 2. MATERYAL VE METOT

### 2.1 Veri Seti

Bu çalışmada kullanılan veri seti tarafından bir proje kapsamında toplanan ve Ekim 2017 tarihi itibarıyla Physionet.org sitesinde yer verilen tuş vuruş datalarıdır [19]. Goldberger ve arkadaşları tarafından gerçekleştirilen proje kapsamında; Amerika, Kanada ve Avusturalya gibi ülkelerden projenin yayınlandığı internet sitesini ziyaret eden kişilerden projeye katılmak isteyenler, günlük Windows uygulamaları (internette arama yapma, mail yazıp gönderme, ve diğer kelime yazma işlemleri gibi...) sırasındaki klavye kullanımlarını kayıt altına almayı sağlayan 'Tappy' isimli uygulamayı bilgisayarlarına indirerek, çalışmaya gönüllü olarak katkıda bulunmuşlardır. Bu proje çalışması İnsan Araştırmaları Etiği Komitesi (Human Research Ethics Committee) tarafından onaylanmış ve yazılımı yüklemeye önce katılımcılar bilgilendirilerek onayları alınmıştır.

Veri seti toplam 216 katılımcının, klavye tuşlarını kullanımlarına ilişkin zaman bilgilerini içermektedir. Katılımcıların 166 tanesi Parkinson hastası, 50 tanesi ise sağlıklı bireylerden oluşmaktadır. Bu bireylere ait toplamda 551 adet kayıt bulunmaktadır. Kayıtlardaki örnek sayıları sabit olmamakla birlikte bazı kayıtlar çok az örnek içermektedir. Hesaplanan özelliklerin örnek sayısına bağımlılığını ortadan kaldırmak için ilk olarak örnek sayısı 1000'in altında olan kayıtlar çalışmaya dahil edilmemiş (bu durumda 515 kayıt üzerinde analizler gerçekleştirilmiştir), ikinci olarak da çok uzun örnekler içeren kayıtlar 1000 civarı örnek içerecek şekilde bölünerek kullanılmıştır (böylece 646 kayıt elde edilmiştir). Örnek azlığı nedeniyle çıkarılan kayıtlar dikkate alındığında katılımcı sayısı; 147 tanesi Parkinson hastası, 49 tanesi de sağlıklı birey olmak üzere 196 olmuştur. Veri setinde yer alan her bir kayıt iki adet dosya içermektedir. Bu dosyalardan biri katılımcılara ait kişisel bilgileri içerirken diğeri katılımcıların klavye kullanımları sırasında elde edilen zaman verilerini içermektedir.

Katılımcıların detaylı kişisel bilgilerinin olduğu klasör Çizelge 2.1'de detaylı şekilde anlatılmıştır. Bu klasördeki veriler 12 adet sütundan oluşmaktadır.

Çizelge 2.1 Katılımcıların detaylı kişisel bilgilerinin olduğu klasörün veri elemanları ve detayları

Veri Elemanları	Detayları
Doğum Yılı (Yaş)	Katılımcının doğum yılını gösterir. 65 yaş ve üstündeki kişilerin %2'sinde bu hastalık görülmektedir.
Cinsiyet	Katılımcının cinsiyeti.
Parkinson (Evet/Hayır)	Katılımcı Parkinson hastası mı değil mi?
Titreme (Tremor) (Evet/Hayır)	Katılımcıda titreme var mı yok mu?
Teşhis Yılı	Katılımcıya eğer hasta ise ne zaman Parkinson hastalığı teşhisi konduğunu ifade eder.
Taraflik (Sol, Sağ, Yok)	Parkinson hastalarında semptomlar tek bir tarafta görülebilir. Eğer bu durum varsa hangi tarafta olduğunu ifade eder.
BPHDÖ (1 ile 5 arası)	1 ile 5 arasında olan bir ölçektir. Hasta tarafından biliniyorsa verilere eklenmiştir. Bu veri setinde sadece 8 adet hastanın BPHDÖ bilgileri bilinmektedir.

Çizelge 2.1 devam ediyor.

Etkisi (Az, Orta, Şiddetli)	Hastalığın şiddeti ve günlük yaşama olan etkisi.
Tedavi için ilaç kullanıp kullanmadığını gösterir. Eğer kullanıyorsa hangi ilaçları aldığını ifade eder.	Parkinson hastalarının ilaç alıp almadığını ifade eder. Levadopa, MAO-B İnhibitörü, Dopamine veya başka bir ilaç olabilir. Özellikle Levadopa motor semptomların karakteristiklerini azaltmaktadır.

Şekil 2.1'de bir katılımcıya ait detaylı kişisel bilgilerinin olduğu dosyanın ekran görüntüsü gösterilmiştir.

```
BirthYear: 1953
Gender: Male
Parkinsons: True
Tremors: False
DiagnosisYear: 2012
Sided: Left
UPDRS: Don't know
Impact: Medium
Levadopa: True
DA: True
MAOB: True
Other: False
```

Şekil 2.1 Bir katılımcıya ait detaylı kişisel bilgilerin olduğu dosyanın ekran görüntüsü

2. klasör ise katılımcıların bilgisayar klavyesi ile olan etkileşimlerinden elde edilen zaman sonuçlarını içermektedir. Bu klasördeki veriler 8 adet sütundan oluşmaktadır. Çizelge 2.2'de detaylı şekilde anlatılmıştır.

Çizelge 2.2 Bir katılımcının bilgisayar klavyesi ile olan etkileşiminden elde edilen dinamikler.

Kayıt İçeriği	Detayları
Katılımcı Tanımlayıcısı	Her bir kullanıcıya 10 karakterden oluşan bir kod isim verilmektedir. Gizlilik politikası için bu şekilde yapılmaktadır.
Tarih	Kaydın hangi tarihte alındığını gösterir (Yıl: Ay: Gün: şeklinde).
Zaman Damgası	Kaydın hangi zaman dilimleri içerisinde gerçekleştiğini gösterir (Saat: Dakika: Saniye: şeklinde).



Çizelge 2.2 devam ediyor.

El	O an kaydedilen kaydın hangi el ile bilgisayar klavyesine basıldığını gösterir (Sol ya da Sağ).
Basılı Tutma Süresi (Hold Time)	O an ki tuşun basılıp serbest bırakılması arasında geçen süredir.
Yön	Bir önceki tuştan bir sonraki basmaya geçişi ifade eder. 8 adet kombinasyon vardır. LL, RR, SS, LS, SL, RL, LR, SR, RS. S Space (Boşluk) tuşunu ifade etmektedir.
Gecikme Süresi (Latency Time)	Bir önceki tuşa bastığı an ile bir sonraki tuşa bastığı an arasında geçen süredir.
Geçiş Süresi (Flight Time)	Bir önceki tuşun serbest bırakıldığı an ile bir sonraki tuşa basıldığı zaman arasında geçen süredir.

Şekil 2.2'de bir katılımcıya ait zaman verilerinin olduğu dosyanın ekran görüntüsü aşağıda gösterilmiştir.

3MZWDTW7CC	161001	16:52:24.289	L	0214.8	LL	0480.5	0273.4
3MZWDTW7CC	161001	16:52:24.703	L	0207.0	LL	0421.9	0207.0
3MZWDTW7CC	161001	16:52:26.312	L	0230.5	RL	0777.3	0640.6
3MZWDTW7CC	161001	16:54:02.594	L	0199.2	LL	0414.1	0136.7
3MZWDTW7CC	161001	16:54:22.719	L	0179.7	LL	0328.1	0097.7
3MZWDTW7CC	161001	16:54:23.074	R	0136.7	LR	0398.4	0218.8
3MZWDTW7CC	161001	16:54:23.340	L	0183.6	RL	0218.8	0082.0
3MZWDTW7CC	161001	16:54:23.641	L	0183.6	LL	0300.8	0117.2
3MZWDTW7CC	161001	16:54:23.984	L	0191.4	LL	0335.9	0152.3
3MZWDTW7CC	161001	16:54:24.223	R	0152.3	LR	0277.3	0085.9
3MZWDTW7CC	161001	16:54:27.398	R	0144.5	LR	0378.9	0179.7
3MZWDTW7CC	161001	16:54:27.617	L	0183.6	RL	0179.7	0035.2
3MZWDTW7CC	161001	16:54:27.785	R	0101.6	LR	0250.0	0066.4
3MZWDTW7CC	161001	16:54:28.133	R	0117.2	RR	0323.0	0230.5
3MZWDTW7CC	161001	16:54:28.562	R	0113.3	RR	0433.6	0316.4
3MZWDTW7CC	161001	16:54:29.027	L	0199.2	RL	0378.9	0265.6
3MZWDTW7CC	161001	16:54:29.312	R	0136.7	LR	0347.7	0148.4
3MZWDTW7CC	161001	16:54:29.785	R	0175.8	RR	0433.6	0296.9
3MZWDTW7CC	161001	16:56:19.141	L	0214.8	LL	0332.0	0136.7
3MZWDTW7CC	161001	16:56:19.523	R	0140.6	LR	0457.0	0242.2
3MZWDTW7CC	161001	16:56:19.820	L	0183.6	RL	0253.9	0113.3
3MZWDTW7CC	161001	16:56:20.188	L	0183.6	LL	0367.2	0183.6
3MZWDTW7CC	161001	16:56:42.730	R	0164.1	LL	0308.6	0109.4
3MZWDTW7CC	161001	16:56:43.109	L	0132.8	LR	0410.2	0246.1
3MZWDTW7CC	161001	16:56:43.375	L	0183.6	RL	0214.8	0082.0
3MZWDTW7CC	161001	16:56:43.680	L	0187.5	LL	0300.8	0117.2
3MZWDTW7CC	161001	16:56:44.062	L	0183.6	LL	0386.7	0199.2
3MZWDTW7CC	161001	16:56:44.262	R	0136.7	LR	0246.1	0062.5
3MZWDTW7CC	161001	16:56:47.285	R	0168.0	LR	0351.6	0168.0
3MZWDTW7CC	161001	16:56:47.492	L	0187.5	RL	0187.5	0019.5
3MZWDTW7CC	161001	16:56:47.676	R	0113.3	LR	0257.8	0070.3
3MZWDTW7CC	161001	16:56:47.996	R	0121.1	RR	0312.5	0199.2
3MZWDTW7CC	161001	16:56:48.414	R	0136.7	RR	0402.3	0281.3
3MZWDTW7CC	161001	16:56:48.816	L	0195.3	RL	0343.8	0207.0
3MZWDTW7CC	161001	16:56:49.012	R	0128.9	LR	0261.7	0066.4
3MZWDTW7CC	161001	16:56:49.398	R	0148.4	RR	0367.2	0238.3
3MZWDTW7CC	161001	16:58:13.375	R	0152.3	RR	0515.6	0359.4
3MZWDTW7CC	161001	16:58:13.914	S	0168.0	RS	0523.4	0371.1
3MZWDTW7CC	161001	16:58:14.316	L	0152.3	SL	0418.0	0250.0
3MZWDTW7CC	161001	16:58:14.812	L	0164.1	LL	0484.4	0332.0
3MZWDTW7CC	161001	16:58:15.125	R	0109.4	LR	0367.2	0203.1
3MZWDTW7CC	161001	16:58:15.527	R	0152.3	RR	0359.4	0250.0

Şekil 2.2 Bir katılımcıya ait zaman ve yön verileri

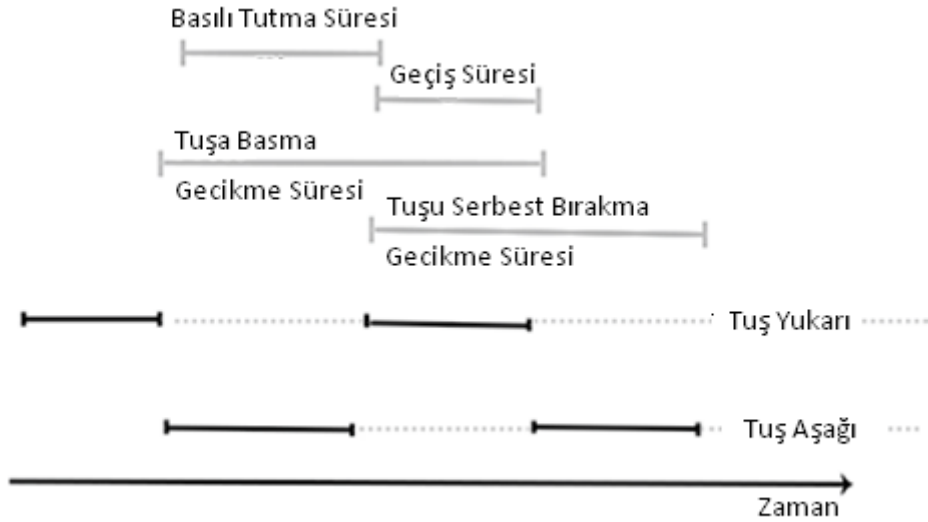
## Tuş Vuruş Dinamikleri

Tuş vuruş dinamikleri, her kişinin kendine has klavyede yazma biçimlerinin olduğunu kabul eden biyometrik bir yöntemdir. Kişilerin klavyelere yapmış oldukları tuş vuruşları arasındaki süre akışının kişi için ayırt edici olabileceği düşünülmektedir. Tuş vuruş dinamiklerinin hesaplanmasında kullanılan bazı nitelikler Şekil 2.3’de detaylı olarak gösterilmiştir.

**Basılı tutma süresi (Hold Time):** Bir tuşa yapılan basma ve çekme arasında geçen zamanı ifade etmektedir.

**Gecikme süresi ( Latency Time):** Bir önceki tuşa yapılan basma ile bir sonraki tuşa yapılan basma arasında geçen süredir.

**Geçiş süresi (Flight Time):** Bir önceki tuştan elin çekilmesi ile bir sonraki tuşa yapılan basma arasında geçen süredir.



Şekil 2.3 Basılı tutma, gecikme ve geçiş sürelerini gösteren diyagram.

## Tuş Vuruş Dinamiklerinin Avantajları

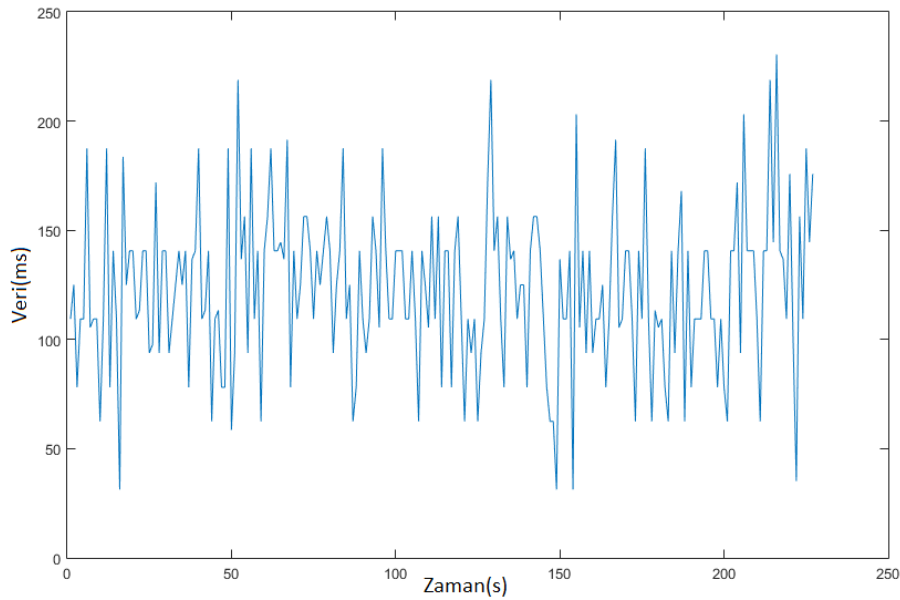
- 1) Tuş vuruş dinamikleri dışarıdan herhangi bir müdahaleye açık değildir.
- 2) Tuş vuruş dinamiklerinin kullanılması çok basittir.
- 3) Tuş vuruş dinamikleri biyometrik kimlik doğrulama ve hastalık tespiti yapan cihazlar kadar pahalı değildir.

- 4) Tuş vuruş dinamiklerini kullanan sistemler ekstra bir donanıma ihtiyaç duymamaktadır.
- 5) Tuş vuruş dinamikleri kullanıcıya ait özel davranışlardır.

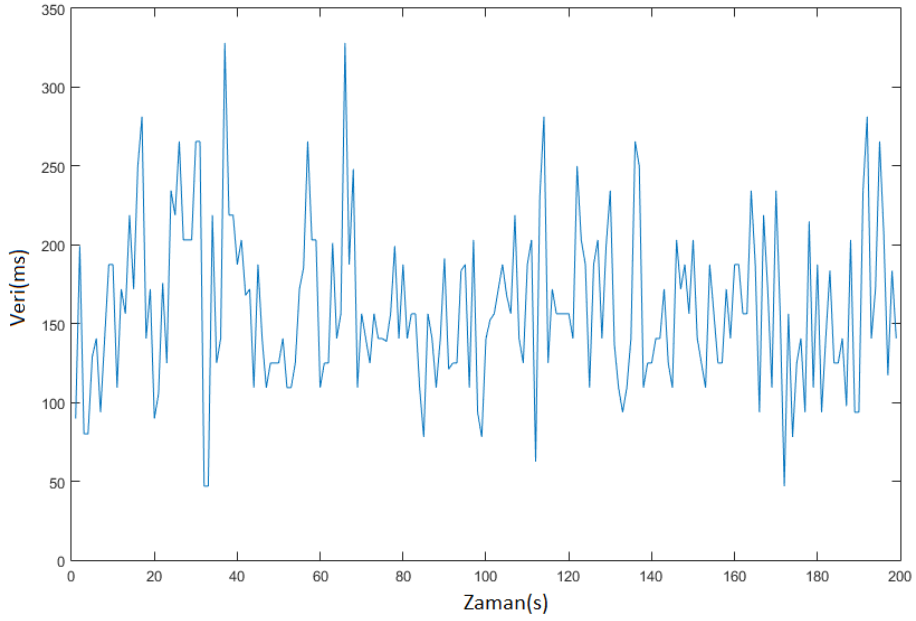
### Tuş Vuruş Dinamiklerinin Dezavantajları

- 1) İnsanların yazma ritimleri o anki hal durumlarına bağlı olarak değişkenlik gösterebilir.
- 2) Elde ve kolda meydana gelen kırık ve yaralanmalar yazma ritminde farklılıklara sebebiyet gösterebilir.
- 3) Klavyenin sertliği, yeniliği, modeli ve tuşların sıralanması gibi faktörler dinamiklerde farklılıklara yol açabilmektedir.
- 4) Yazma stili, yazılan metnin diline bağlı olarak değişiklik gösterebilir. Anadilde yazılan bir metin ile yabancı dilde yazılan bir metnin yazılma hızı farklılık gösterebilir.

Parkinson hastası ve sağlıklı bir katılımcıya ait basılı tutma sinyallerinin örnekleri Şekil 2.4 ve Şekil 2.5'de verilmiştir.



Şekil 2.4 Parkinson hastası bir katılımcının basılı tutma verisi.



Şekil 2.5 Sağlıklı bir katılımcının basılı tutma verisi.

## 2.2 Çıkarılan Özellikler

### 2.2.1 Asimetri hesabı ( Simetri Bozukluğu)

Parkinson hastalığı ilk başta insanın tek bir tarafında başlayabilirken bazen her iki tarafında da meydana gelebilmektedir. Bu çalışmada kullanılan verilerde kişinin hangi anda hangi tuşa bastığı bilinmemektedir ancak kişinin o anda bastığı tuşa hangi eliyle bastığı bilinmektedir. Bu yüzden çalışmamızda soldan sola ve sağdan sağa geçişler ile soldan sağa ve soldan sağa geçişler arasındaki farklılıklar incelenmiş ve simetri bozukluğu olup olmadığı anlaşılmaya çalışılmıştır. Her geçişte yeni bir tuşa basılmıştır. Bu yaklaşıma göre yapılan işlem Eşitlik 2.1 ve Eşitlik 2.2'de gösterilmektedir.

$$f_{asy}(LL - RR) = \left| \frac{1}{a} \sum_{i=1}^a (LL)_i - \frac{1}{b} \sum_{j=1}^b (RR)_j \right| \quad (2.1)$$

Eşitlik 2.1'de yer alan LL; soldan sola geçişleri, RR sağdan sağa geçişleri, a; verideki soldan sola geçişlerin sayısını ve b; verideki sağdan sağa geçişlerin sayısını ifade etmektedir.

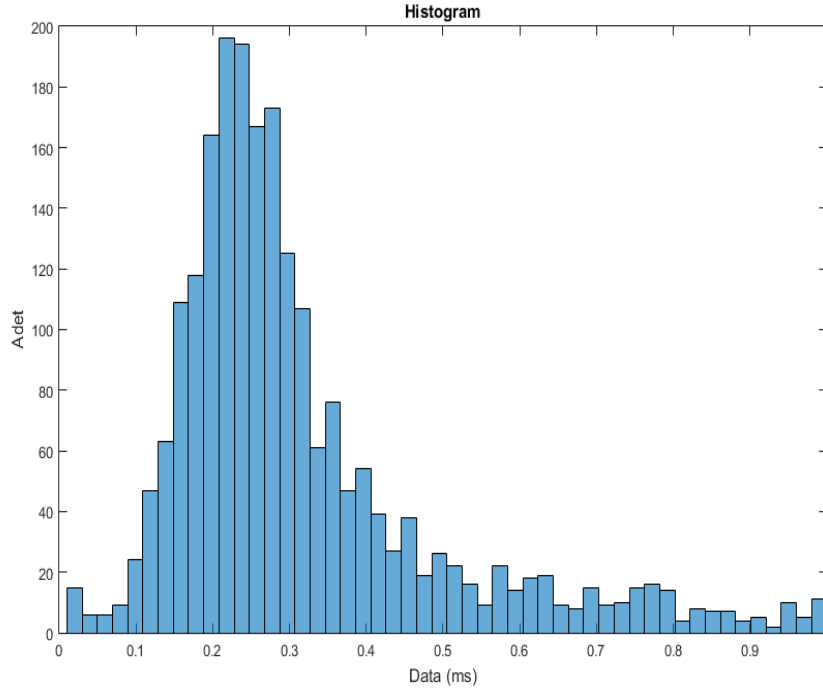
$$f_{asy}(LR - RL) = \left| \frac{1}{c} \sum_{i=1}^c (LR)_i - \frac{1}{d} \sum_{j=1}^d (RL)_j \right| \quad (2.2)$$

Eşitlik 2.1'de yer alan LR; soldan sağa geçişleri, RL sağdan sola geçişleri, c; verideki soldan sağa geçişlerin sayısını ve b; verideki sağdan sola geçişlerin sayısını ifade etmektedir.

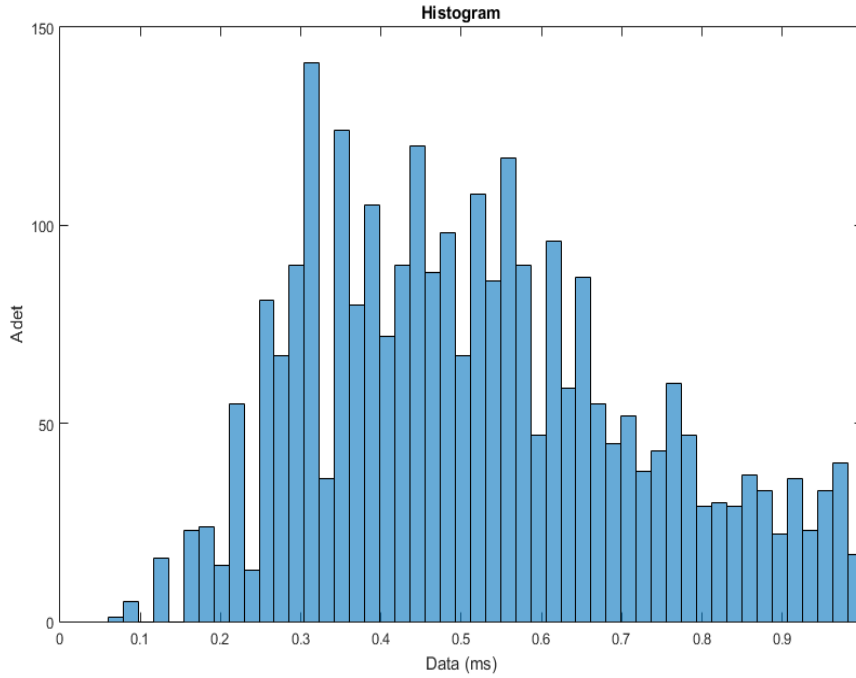
### 2.2.2 Histogram alanları

Histogram, bir zaman serisindeki veri dağılımını göstermektedir. Yani tekrarlı sayılardan oluşan veriyi önce tabloya sonra tabloyu grafiğe dönüştürmektedir. Elde edilen grafik, bir zaman serisine ait verilerin hangisinden kaç tane içermekte olduğunu, dikdörtgen sütunlar halinde göstermektedir.

Yapılan bu çalışmada basılı tutma, gecikme ve geçiş sürelerinin ayrı ayrı histogramları elde edilmiştir. Katılımcıya ait zaman serisi 50 eşit kutucuğa bölünmüştür. Basılı tutma, gecikme ve geçiş sürelerinin maksimum ve minimum değerleri her bir katılımcıda farklı olduğu için normalizasyon işlemi uygulanmış ve x eksenini sıfır bir arasında normalize edilmiştir. Elde edilen histogramda her bir kutucuğun alanı ayrı ayrı hesaplanmıştır. Alan hesabı, her bir kutucuğun gösterdiği süreler o zaman serisinde bulunduğu miktar ile çarpılarak yapılmıştır. Katılımcıların her birine uygulanan kayıt sayısı farklı olduğu için elde edilen alan hesaplamalarına da tekrar normalizasyon işlemi uygulanmıştır. Bu işlem her bir kutucuktaki alanın toplam alana bölünmesiyle yapılmıştır. Şekil 2.7 ve 2.8'de Parkinson hastası ve sağlıklı bir bireyin histogramları gösterilmiştir.



Şekil 2.6 Parkinson hastası bir katılımcının basılı tutma sürelerinin (Hold Time) histogram'ı.



Şekil 2.7 Sağlıklı bir katılımcının basılı tutma sürelerinin (Hold Time) histogram'ı.

Şekil 2.6 ve 2.7'de görüldüğü üzere x eksenı (basılı tutma süreleri) 0-1 arasında normalize edilmiştir. Y eksenı (o kayıтта kaç adet basılı tutma kaydedildiđi) şekil

üzerinde normalize edilmemiştir ancak her bir bölgenin alan hesabı yapılırken o bölgedeki alan toplam alana bölünerek normalizasyon işlemi gerçekleştirilmiştir.

### **2.2.3 Ortalama, Medyan, Standart Sapma, Varyans Çarpıklık ve Basıklık parametreleri**

Basıklık ve çarpıklık değerleri verilerin normal dağılım gösterip göstermediğini ifade eder. Ortalama 1.moment; varyans 2. moment olarak bilinirken çarpıklık 3. ve basıklık ise 4. moment olarak bilinmektedir. Bu özelliklerin olasılıksal dağılımları istatistiksel fonksiyonlar ile hesaplanmaktadır [20, 21, 22].

Çarpıklık veri dağılımının normalden uzaklaşarak ne kadar sağa veya sola doğru yönelerek yamuk bir şekil aldığını ifade eden bir kavramdır. Normal bir dağılımda çarpıklık sıfırdır. Çarpıklık katsayısı  $-\infty$  ile  $+\infty$  arasında değerler alabilmektedir. Bu yüzden pozitif ve negatif çarpıklık olmak üzere iki tür çarpıklıktan söz edilebilir. Bazı durumlarda tanımsız da olabilir. Eğer ortalama medyandan küçük ise dağılım sola (negatif) çarpık olur. Eğer ortalama medyandan büyükse dağılım sağa (pozitif) çarpık olur. Çarpıklık değerleri -3 ile +3 arasında değer almış sa normal dağılım olarak kabul edilmektedir.

Basıklık ise normal dağılım eğrisinin ne kadar dik veya basık olduğunu gösterir. Tam çan eğrisinin basıklık katsayısı sıfırdır. Basıklık katsayısı pozitif ise eğri normale göre daha dik olur. Basıklık katsayısı negatif ise eğri daha basık olur. +1.96 ile - 1.96 arasındaki değerlerde eğrinin dik olmadığı söylenebilir.

Ortalama hesabı Eşitlik 2.3'te gösterilmiştir.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.3)$$

Eşitlik 2.3'te yer alan  $\bar{x}$  örneklem aritmetik ortalaması sembolüdür; n ise veri uzunluğudur.

Medyan hesabı Eşitlik 2.4'de gösterilmiştir.

$$\text{Medyan} = \begin{cases} \left(\frac{n+1}{2}\right). \text{terim} & ; n \text{ tek ise} \\ \frac{\left(\frac{n}{2}\right). \text{terim} + \left(\frac{n}{2} + 1\right). \text{terim}}{2} & ; n \text{ çift ise} \end{cases} \quad (2.4)$$

Varyans hesabı Eşitlik 2.5'de gösterilmiştir.

$$\sigma^2(F) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.5)$$

Standart Sapma hesabı Eşitlik 2.6'da gösterilmiştir.

$$\sigma = \sqrt{\text{Var}(X)} \quad (2.6)$$

Çarpıklık Katsayısı (ÇK) hesabı Eşitlik 2.7'de gösterilmiştir.

$$\text{ÇK} = \frac{\sqrt{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\sigma^3} \quad (2.7)$$

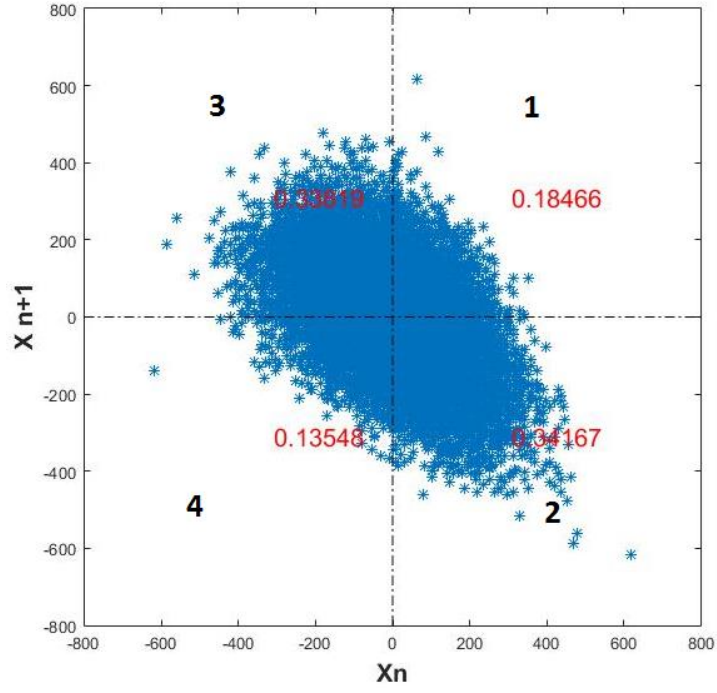
Basıklık Katsayısı (BK) hesabı Eşitlik 2.8'de gösterilmiştir.

$$\text{BK} = \frac{\sqrt{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\sigma^4} - 3 \quad (2.8)$$

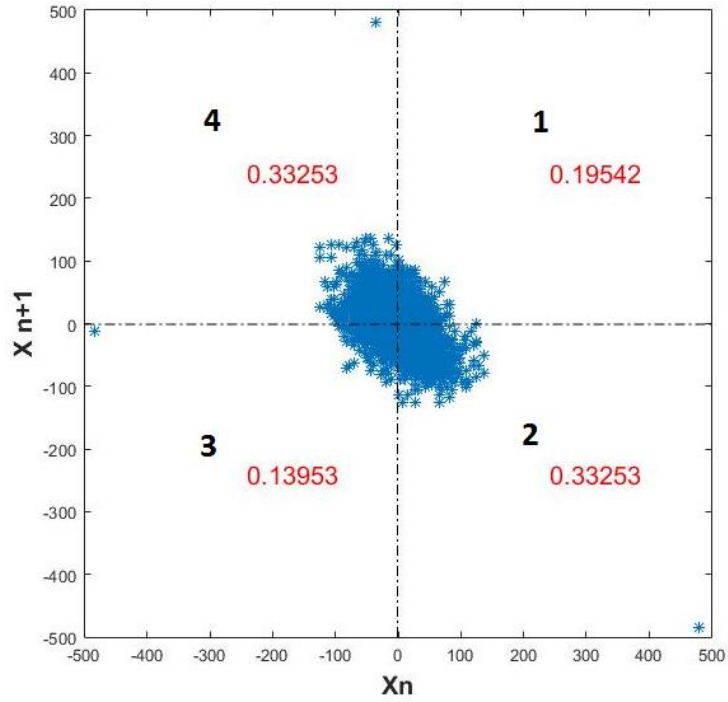
#### 2.2.4. Durum uzayı ( $X_n$ vs $X_{n+1}$ )

Kullanılan zaman serisi =  $x_n$  olsun.  $x_n$ 'in  $x_{(n+1)}$ 'e karşı çizdirilmesiyle durum uzayı elde edilmektedir. Yatay eksen  $x_n$ 'i, düşey eksen ise  $x_{(n+1)}$ 'i göstermektedir. Yani zaman serisinin bir sonraki an ile mevcut an arasındaki benzerliği incelenmektedir. Bu çalışmada durum uzayı Sağ Üst (1. Bölge), Sağ Alt (2. Bölge), Sol Üst (3. Bölge) ve Sol Alt (4. Bölge) olmak üzere dört eşit bölgeye ayrılarak, durum uzayındaki noktaların bu bölgeleri ziyaret etme sayılarının normalize değerleri hesaplanmıştır. Normalizasyon işlemi için toplam nokta sayısı kullanılmıştır. Şekil 2.9 ve 2.10'da sırasıyla Parkinson hastası ve sağlıklı bir katılımcıya ait durum uzayları ve her bölge için elde edilen normalize değerler gösterilmektedir.





Şekil 2.8 Parkinson hastası bir katılımcıdan elde edilen durum uzayı.



Şekil 2.9 Sağlıklı bir katılımcıdan elde edilen durum uzayı.

## 2.2.5. İlk fark (First Differencing) alınarak elde edilen durum uzayı

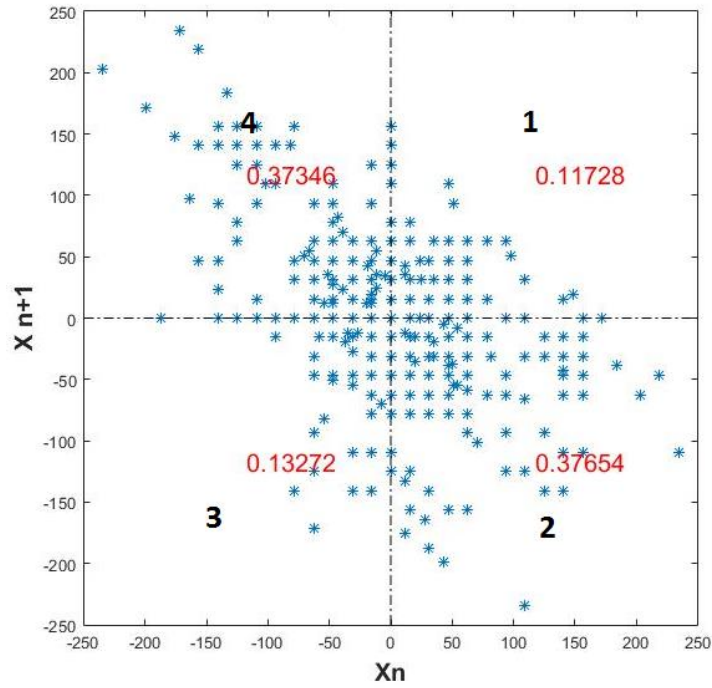
Zaman serisi içerisindeki ardışık iki veri arasındaki farklar elde edilerek oluşturulan yeni bir zaman serisi içinde durum uzayı oluşturulmuştur. Oluşturulan yeni seride yer alan değerler Eşitlik 2.9'daki gibi tanımlanabilir.

$$Dif_{(n)} = x_{(n+1)} - X_{(n)} \quad (2.9)$$

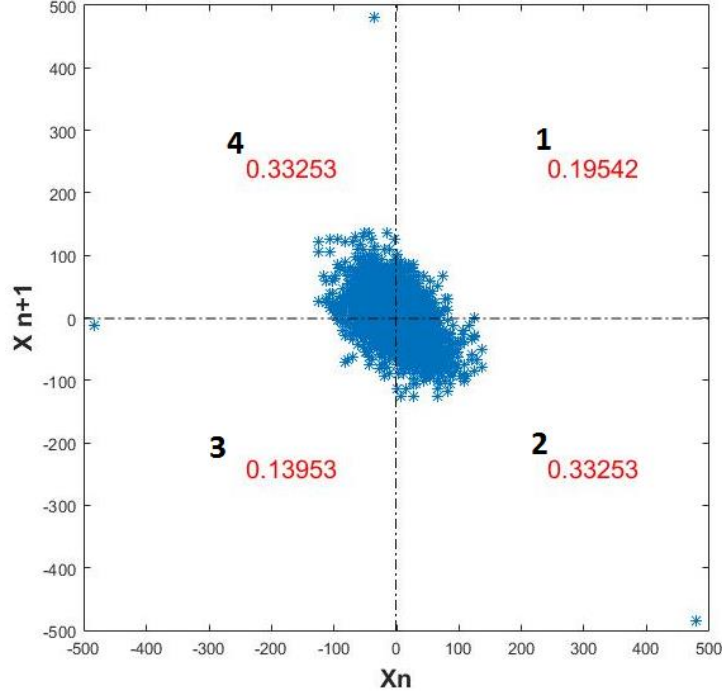
Burada  $X$  kullanılan zaman serisini,  $n$  ise o zaman serisindeki indisi ifade etmektedir.

$Dif_{(n)}$ 'in  $Dif_{(n+1)}$ 'e karşı çizdirilmesiyle ilk fark durum uzayı elde edilmektedir.

Şekil 2.10 ve 2.11'de sırasıyla Parkinson hastası ve sağlıklı bir katılıcı için oluşturulan ilk fark zaman serisinde elde edilen durum uzayı yer almaktadır.



Şekil 2.10 Parkinson hastası bir katılımcıdan elde edilen ilk fark durum uzayı.



Şekil 2.11 Sağlıklı bir katılımcıdan elde edilen ilk fark durum uzayı.

## 2.2.6 Entropi

Entropi, fizikte bir sistemin mekanik işe çevrilemeyecek termal enerjisini temsil eden termodinamik terimidir [23]. Genellikle bir sistemdeki düzensizlik ve rasgelelik olarak tanımlanır. İstatistikten teolojiye, felsefeye kadar pek çok alanda kullanılır. Sembolü  $S$ ' dir ve termodinamiğin ikinci yasası olarak bilinir. Bu yasa, hiçbir cihazın veya sistemin aldığı ısıyı tamamen işe dönüştüreceği şekilde çalışamayacağını, ayrıca sadece ısıyı bir sıcaklıktan daha yüksek bir sıcaklığa transfer eden bir sürecin olmadığını açıklamaktadır. Buna ek olarak ne olursa olsun, enerjinin ve kütlenin olduğu her yerde entropi mevcut olduğunu açıklamaktadır.

Fen bilimlerinin en önemli yasası her şeyin yıprandığını söyleyen yasadır. Canlılar yaşlanır ve ölürlere, arabalar paslanır ve evrendeki düzensizlik artar. Bilim adamları düzensizliği Entropi denilen nicelik ile ölçerler. Sistemlerdeki düzensizlik arttıkça, entropi de artar. Bu durum da faydalı (iş yapılabilir) enerji miktarını azaltır, faydasız enerjiyi (entropi) artırır.

Klasik termodinamikte sıcaklık, basınç, hacim, enerji ve entropi gibi kavramlar temel alınır. Bunun haricinde termodinamik istatistiksel kavramlar kullanılarak da ifade edilebilir.

Bilgi kuramında entropi bir iletinin bilgi içeriğini ölçer. Yani, ölçülen entropi ile sistemin kişiye ne kadar yeni bilgi katacağını ifade eder. Bu bağlamda entropi ilk defa 1948 yılında Claude E. Shannon tarafından ortaya atılmıştır. Ayrık bir rassal değişkenin entropisi Eşitlik 2.10'da gösterilmiştir.

$$H(X) = - \sum_i p(x_i) \log p(x_i) \quad (2.10)$$

Eşitlik 2.10'da yer alan  $p$  olasılık fonksiyonunu ifade etmektedir.

Entropinin istatistik biliminde de ayrı bir tanımı vardır. Ludwig Boltzmann'ın denkleminde entropi,  $S$ , bir sistemin girebileceği mikroskopik durumların sayısı,  $W$  yoluyla tanımlanır. Eşitlik 2.11'de gösterilmiştir. Burada  $k$  Boltzmann sabitidir. Sözü edilen mikroskopik durumların tanımı ve sayılması ise, sistemi oluşturan atomları tarif eden mekanik yasalar kullanılarak yapılır.

$$S = k \cdot \log W \quad (2.11)$$

Bu çalışma kapsamında Shannon Entropi, Örnek Entropi ve Yaklaşık Entropi kullanıldı.

### **Shannon Entropi Hesaplanması**

İsmi bilgi teorisinin babası olan Claude Shannon'dan almaktadır. Shannon buradaki  $H$  ismini Ludwig Boltzmann'ın termodinamikteki  $H$ -teoremine atfen seçmiştir. Orijini Pauli ve Von Neumann'a kadar dayanmaktadır. Bilgi teorisindeki temel problem büyük miktardaki bilgilerin kodlanmasıyla baş etmeye çalışmasıdır. Sonlu bir veri seti ( $A$ ) ile başlanır, bu set Latin alfabesindeki 26 harfle, ASCII kodundaki 128 sembol ile veya daha geniş kelime setleri ile gösterilebilir. Bilginin kayıpsız bir şekilde kodlanabilmesi için kaç bite ihtiyaç duyulduğu cevap aranan sorudur [24].

$A$ 'nın Shannon entropi hesabı Eşitlik 2.12'de gösterilmiştir.

$$H(A) = - \sum_{i=1}^n P_i \log_2 P_i \quad (2.12)$$

Eşitlik 2.12'de yer alan  $P_i$ ; olasılık kütle fonksiyonunu ifade etmektedir.

### **Yaklaşık Entropi (YakEn) (Approximate Entropy)**

İstatistikte kullanılan bu teknik bir zaman serisi içerisindeki dalgalanmaların öngörülemezliğinin ve düzenliliğinin miktarını ölçmeye yarar [25]. Genellikle 100 kayıttan fazla ve gürültülü zaman serilerinde kullanılan bir metottür. Matematiksel olarak, Yaklaşık Entropi genel bir yaklaşım olup Markov zincirinin bir sürece yaklaştırılmasıyla oluşan entropi değeridir. Çıkan sonuç çok küçük bir değerse (Örnek: 0.00001) veri dizisi düzenli ve tahmin edilebilirdir anlamına gelmektedir. Tekrarlanan dalgalanma modellerinin olduğu bir zaman serisi, tekrarlanmaların olmadığı bir zaman serisine göre daha öngörülebilirdir. YakEn gözlemlerin benzer modellerinin ek olarak benzer modeller ile takip edilemeyeceğinin olasılığını yansıtır. Eğer bir zaman serisi çok miktarda tekrar eden benzer modelleri içermekteyse YakEn değeri göreceli olarak daha düşük olacaktır. Tahmin edilebilirliği az ise YakEn değeri artacaktır.

$$\varphi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log(C_i^m(r)) \quad (2.13)$$

$$\text{YakEn} = \varphi^m(r) - \varphi^{m+1}(r) \quad (2.14)$$

$\varphi^m(r)$ ; tüm model içirişinden alınan bir pencerenin (kesitin) modelin tümüne ortalama ne kadar benzediğini ölçer.

Eşitlik 2.13 ve 2.14'de yer alan;  $m$ , lag'dır. Genellikle 2 veya 3 olarak kabul edilir ancak uygulamaya göre farklılık gösterebilir. Yapılan bu çalışmada her bir zaman serisi için ortalama ortak bilgi fonksiyonu hesaplatıldı ve fonksiyonun ilk minimum olduğu yer  $m$  (lag) olarak alınmıştır.  $r$  tolerans değeridir ve bu değer 0.1 ile 0.25 arasında olabilir. Bu tez çalışmasında  $r$  değeri 0.2 olarak kullanılmıştır.  $N$  ise girişteki veri noktalarının sayısıdır. Genellikle 75 ile 5000 arasında olması istenir.

### **Örnek Entropi (ÖrnEn) (Sample Entropy)**

Örnek entropi (ÖrnEn), yaklaşık entropi'nin (YakEn) değiştirilmiş halidir. Tıpkı YakEn gibi zaman serilerinin karmaşıklığını ölçmektedir. YakEn'de olduğu gibi ÖrnEn de 3 parametreden oluşmaktadır ( $m, r, N$ ; bu parametreler YakEn başlığı altında açıklanmıştır). Fizyolojik bir zaman serisi sinyalinin kompleksliğinin

değerlendirilmesinde ve hastalık durumlarının tespit edilmesinde kullanılır [26]. ÖrnEn'in YakEn'e göre iki tane avantajı vardır. Veri uzunluğunun bağımsız oluşu ve nispeten daha kolay uygulanması bu metodu avantajlı kılar. Buna ek olarak hesaplamada küçük bir farklılık vardır. YakEn'de kendiyile olan benzerliğine bakılmaktadır. Bu da olasılıkların ( $C_i^m(r)$ ) asla sıfır olmamasını sağlamaktadır. Bu kendiyile yapılan kıyaslar sonucu YakEn değerleri düşmektedir ve sinyaller gerçekte olduğundan daha düzenli olarak yorumlanmaktadır. ÖrnEn'de bu kendi içerisinde yapılan eşleşmeler yoktur. ÖrnEn değeri mutlaka sıfır veya pozitif olmalıdır. ÖrnEn değerinin küçük olduğu durumlarda kendine benzerlik (self-similarity) çok, gürültü ise az olur. Matematiksel formülü eşitlik 2.15'de gösterilmiştir.

$$\text{ÖrnEn} = -\log A/B \quad (2.15)$$

Bu eşitlikte A; mesafe fonksiyonunun r değerinden küçük olduğu durumdaki m+1 uzunluğunun sahip olduğu taslak vektör çifti sayısını, B; mesafe fonksiyonunun r değerinden küçük olduğu durumdaki m uzunluğunun sahip olduğu taslak vektör çifti sayısını, m; Veri noktalarının uzunluğunu ifade etmektedir. A değeri genellikle B'den küçük veya B'ye eşit olur. Mesafe Fonksiyonu ( $d[X_m(i), X_m(j)] \dots$ ) ile iki vektör arasındaki mesafe ölçülmektedir.

### 2.2.7 Ortalama Ortak Bilgi (Average Mutual Information)

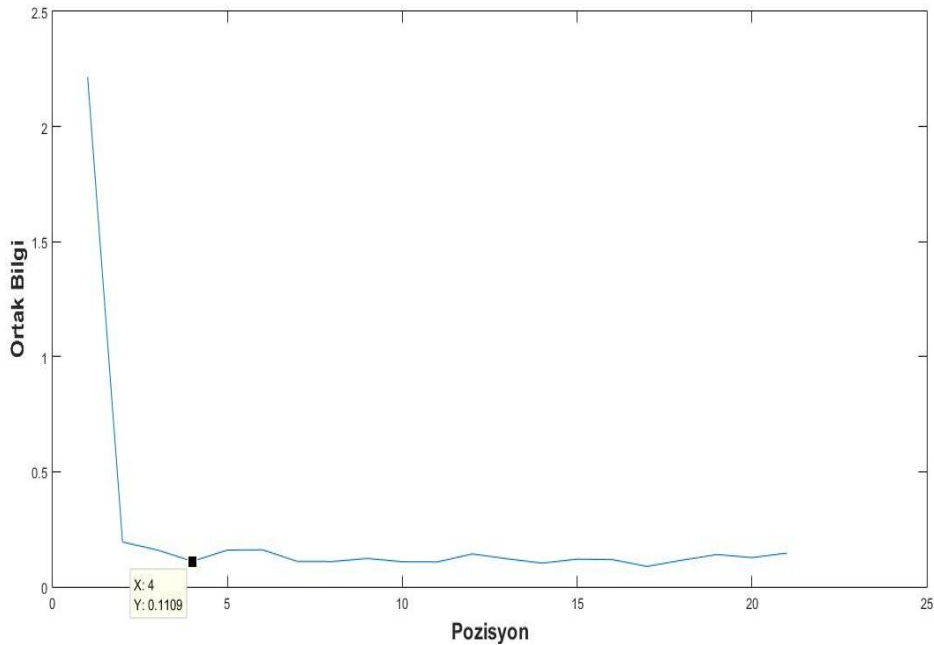
Ortalama ortak bilgi, olasılık teorisinde ve bilgi teorisinde rasgele iki değişkenin aralarındaki ortak bağımlılık derecesini ölçen bir ölçüttür. Daha spesifik olarak; rasgele bir değişkeni gözlemleyerek diğer bir rasgele değişken hakkında ne kadar bilgi elde edilebileceğini ifade eder. Ortak bilgi kavramı, bilgi teorisinde rasgele bir değişkenin içerisindeki entropiye bağlıdır ve o değişkenden elde edilmesi beklenen bilgi miktarını ifade eder [26, 27, 28]. Fraser ve Swinney tarafından ortaya atılmıştır. Temelinde yatan fikir basittir. Tek değişkenli bir sistem olan  $X(t)$ 'yi gözlemleyerek daha yüksek boyutlu bir çeker inşa etmek isteniyorsa, n+1 boyutlu çizilebilen bir koordinat üretmek için ( $X(t), X(t+T), \dots, X(t+nT)$ ) şeklinde düşünülmelidir. Burada T değerini doğru seçmek önemlidir çünkü bu değer gecikmeyi temsil eder. Eğer gecikme çok küçükse  $X(t)$   $X(t+T)$ 'ye çok benzer anlamını taşır ve sistemin çekeri çizildiğinde sınırlar neredeyse üst üste gelmektedir. Bu da kazanılacak bilgi miktarının fazla olduğu anlamına gelir. Eğer gecikme çok fazla ise koordinatlar

birbirlerinden çok bağımsızdır yani benzerlik çok azdır anlamını taşır ve buradan bilgi çıkarılamayacağı anlamına gelir. Buradaki temel fikir şudur: T için iyi bir seçim yapmak, verilen  $X(t)$ ' nin  $X(t+T)$ ' nin ölçülmesiyle ne kadar yeni bilgi katacağıdır.  $X(t)$ ' nin bir ölçüsü verildiğinde  $X(t+T)$  hakkında ortalama kaç bit tahmin edilir sorusunun cevabı Ortak Bilgi ( $I$ ) 'dir. Matematiksel formülasyonu Eşitlik 2.16'da verilmektedir. Bu değerin küçük olması istenilir. Bu değer seçimini yapmak için birden çok metot mevcuttur ancak Fraser ve Swinney ilk minimum noktasını seçmeyi tercih etmişlerdir. Bu çalışmada da ilk minimum noktası ortak bilgi olarak kabul edilmiştir.

$$I(X:Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)} \quad (2.16)$$

Eşitlik 2.16'da X ve Y iki farklı rasgele değişken; P(x,y) ise bunların ortak olasılık fonksiyonudur.

Şekil 2.12'de bir katılımcıya ait  $I(T)$  fonksiyonu yer almaktadır. Fonksiyonun ilk minimum (first minimum) noktası esas alınmıştır ki bu nokta gecikme parametresi T olarak kabul edilir.



Şekil 2.12 Bir katılımcıya ait  $I(T)$  fonksiyonu grafiği.

## 2.2.8 Yerleştirme Boyutu (Embedding Dimension)

Yerleştirme boyutu sisteme ait çekerin durum uzayına yerleştirilmesi / gömülmesi için gerekli olan en düşük boyuta denir [29]. Çeker, sistemin dinamikleri ile ilgili bilgi veren kalıcı durum koşullarıdır. Diğer bir deyişle, ölçümlerden başlayarak ve yörüngenin kendi kendine çakışmadığı bir durum uzayından yeniden inşa edilebilecek bir uzayın minimum boyutudur ve en az sistemin çekerinin boyutu kadar olmalıdır. Sistemin kaç boyuta yerleştirilebileceği anlaşılınca, o sistemin kaç farklı değişkenden etkilendiği de anlaşılmaktadır. Bu da determinizmi doğrulamaktadır. Evrensel bir model yaratılmaya çalışıldığında model minimum boyuta ihtiyaç duymalıdır. Yerleştirme boyutunun belirlenmesinde en yaygın olarak kullanılan metot En Yakın Yanlış Komşu (EYYK) (False Nearest Neighbor-FNN) metodudur. Kullanılabilecek diğer metotlar Bad Prediction Method ve Fractal and Correlation Dimension'dır.

### En Yakın Yanlış Komşu (EYYK) (False Nearest Neighbor - FNN)

En yakın yanlış komşu algoritması yerleştirme boyutunu seçmek için kullanılır. Bu metot Kennel tarafından ortaya atılmıştır [30]. Temelinde yatan fikir, bir sinyal yörüngesindeki (trajectory) komşuların sayısının artan yerleştirme boyutu ile nasıl değiştiğini incelemektir. Düşük yerleştirme boyutunda komşuların çoğu yanlış olur ancak boyut değeri büyütüldükçe komşular gerçek ve doğru olmaktadır. Boyutun artması ile birlikte yanlış olan komşular komşu olmaktan çıkmaktadırlar. Bu nedenle komşuların sayısının bir boyut fonksiyonu olarak nasıl değiştiğini incelemekle, uygun bir yerleştirme boyutu seçilebilir. Algoritma şöyle açıklanabilir; gözlemlenen zaman serisinin gecikmeli koordinatları alınarak m boyutlu durum uzayı inşa edilir. Bu inşa gerçekleşirken gecikme parametresi, ortak bilgi fonksiyonun ilk minimum noktası olarak seçilir. Ardından bu algoritma; m boyutlu uzay içerisindeki her noktanın en yakın komşusu ile olan mesafesini hesaplar. Aynı olay m+1 boyut için tekrar edilir. Eğer Eşitlik 2.17'de hesaplanan değer, önceden belirlenen eşik değerinden büyükse komşular yanlıştır anlamını taşır. Bu olay yanlış komşuların sayısı sıfır olana kadar tekrarlanır.

$$\sqrt{\frac{R^2(m+1) - R^2(m)}{R^2(m)}} \quad (2.17)$$



Eşitlik 2.17'de  $R(m)$ ;  $m$  boyutlu uzay içerisindeki noktaların en yakın komşularına olan mesafesini;  $R(m + 1)$  ise  $m+1$  boyutlu portre içerisindeki noktaların en yakın komşularına olan mesafesini ifade etmektedir.

### 2.3 Özellik Seçimi

Özellik seçimi makine dili öğrenme algoritmalarında ve istatistikte sıkça kullanılmaktadır [31]. Değişken seçimi; nitelik seçimi veya değişken alt küme seçimi olarak da bilinmektedir. Bu olay; ilgili özelliklerin bir alt kümesinin seçilerek yeni bir model inşa edilmesi süreci olarak tanımlanır. Özellik seçimi tekniği 4 farklı sebepten dolayı kullanılmaktadır.

1. Eğitim (Training) süresini kısaltmak için.
2. Boyutsallık sorununu önlemek için.
3. Sistemin ezbere kaçmasını önlemek için.
4. Modeli basitleştirerek kullanıcının veya çalışanın anlamasını kolaylaştırmak için.

Özellik seçiminin esas kullanılma nedeni hangi özelliklerin gereksiz veya konuyla ilgisiz olduğunu anlamak ve bu nedenle oluşacak bilgi ve zaman kaybını ortadan kaldırmak için kullanılır. Veriler üzerinde yapılan çalışmada, tahmini modelleme problemi ile en alakalı özelliklerin otomatik olarak seçilmektedir. Geri beslemeli ve ileri beslemeli olmak üzere 2 tür özellik seçimi mevcuttur. Geri beslemeli özellik seçiminde en gereksiz özellik tüm özelliklerin içerisinde çıkarılarak sırayla devam eder. İleri beslemeli özellik seçiminde ise en gerekli özelliğin üzerine sıradaki en gerekli özellik eklenerek sırayla devam eder.

Yapılan bu çalışmada özellik seçimi için Rastgele Orman(RO) (Random Forest) algoritması ve istatistiksel testler kullanılmıştır.

#### 2.3.1 Varyans Analizi (ANOVA)

Varyans Analizi veya ANOVA, İngilizce Analysis of Variance sözcüklerinin kısaltılmasından gelmektedir [32]. İngiliz genetikçi ve istatistikçi Ronald Fisher tarafından ortaya atılmıştır. Bazı zamanlarda bu analize Fisher'ın varyans analizi de denildiği görülmüştür. İstatistik bilim dalında, grup ortalamaları ve (gruplar içi ve

gruplar arası varyasyon gibi) bunlara bağılı olan işlemleri analiz etmek için kullanılan bir istatistiksel modeldir. Bir deneydeki veya bir anketteki sonuçların ayırt edici olup olmadığını anlamak için kullanılır. Varyans analizi kullanılmaktayken belirlenmiş bir değişkenin gözlemlenen varyansı farklı değişim kaynaklarına dayandırılabilen varyans bileşenine ayrılır. En basit ifadeyle "Varyans Analizi" birden çok grubun ortalamalarının birbirine eşit olup olmadığını denemek için kullanılan çıkartımsal bir istatistik modelidir ve bu deneme iki grup için yapılan t-test denemesini çoklu gruplar için genelleştirir. ANOVA test gruplarını eşit varyanslı olarak kabul eder. Eşit olmayan varyans için bağımsız t- test kullanılır.

Araştırmacılar ve öğrenciler ANOVA'yı farklı yollarla kullanmaktadırlar. 3 farklı şekilde kullanılabilir. Tek Yönlü ANOVA'da bağımsız değişken sayısı birdir. İki Yönlü ANOVA'da ise bağımsız değişken sayısı ikidir. İki bağımsız değişken arasındaki etkileşimi belirlemek için kullanılır. Bu etkileşimler bağımsız değişkenlerin tüm kategorilerinde tek tip olmadığını göstermektedir. N Yönlü ANOVA'da ise bağımsız değişken sayısı ikiden fazla olmaktadır. N değeri bağımsız değişken sayısına göre değişmektedir.

Bu çalışmada varyansın eşit olduğu gruplar arası karşılaştırmalar tek yönlü ANOVA, eşit olmayan varyansa sahip gruplar arası karşılaştırmalarda ise bağımsız t-testi kullanılmıştır.

### **2.3.2 Rastgele Orman (RO) (Random Forest- RF) Algoritması**

Bir özellik seçme algoritması olan Rastgele Orman algoritması, sınıflandırma işlemi sırasında birden çok karar ağacı üreterek sınıflayıcı doğruluk yüzdesini yükseltmeyi hedefler [33]. 2001 yılında Leo Breiman tarafından geliştirilmiştir. Astronomi, sağlık, yazılım ve kontrol sistemleri gibi pek çok alanda kullanılmaktadır. Bazı özellikleri şu şekildedir: Sistemi gürültülü verilerden arındırır, çok kısa sürede sonuç verir, daha iyi genellemeler sunar ve geçerli tahminlerde bulunur. Bu algoritma kullanıcı tarafından iki parametre alır. M; her bir düğümde kullanılan değişkenlerin sayısını, N ise geliştirilecek ağaç sayısını ifade etmektedir.

Eğitim veri setinin 2/3'ünden önyükleme örnekleri oluşturulur. 1/3'lük geri kalan kısım ise hataları test etmek için kullanılır. Her bir düğümdeki m değişkeni tüm değişkenler

arasından rastgele seçilir ve ona göre en iyi dal belirlenir. Optimum sonuca en yakın sonuç, değişken sayısı  $m$ 'nin karekökü olarak alınır.

Sınıfların homojenliğinin ölçülmesi için GINI indeksi kullanılır. Bu indeks değeri ne kadar düşük olursa sınıf o kadar homojendir anlamını taşır. GINI indeksi Eşitlik 2.18'de verilen eşitlik ile hesaplanır.

$$Gini(T) = 1 - \sum_{j=1}^n (P_j)^2 \quad (2.18)$$

Eşitlik 2.18'de  $T$ ; tüm veri setini,  $P_j$ ; veri setindeki her bir verinin, kendisinden küçük ve kendisinden büyük eleman sayılarına bölüm karesi,  $n$  ise seçilen veriyi ifade etmektedir.

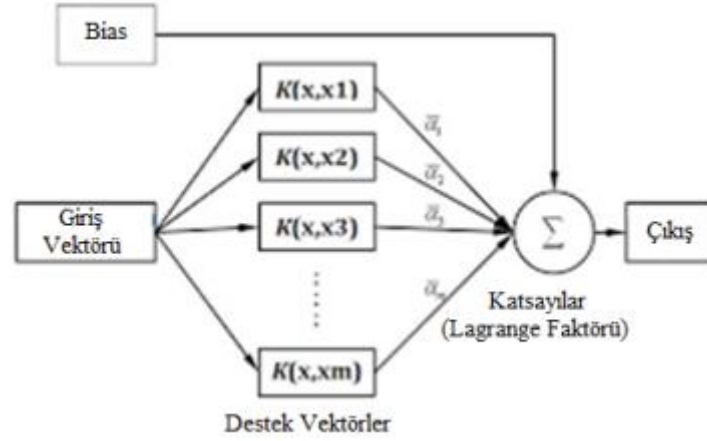
## 2.4 SINIFLAMA ALGORİTMALARI

Bu tez çalışmasında, elde edilen tüm özellikler Destek Vektör Makinesi ve  $k$  En Yakın Komşu sınıflayıcıları ile sınıflandırılmıştır.

### 2.4.1 Destek Vektör Makinesi (DVM) (Support Vector Machines-SVM's)

Destek Vektör Makineleri (DVM), yapısal risk minimizasyonu prensibine göre çalışan dış bükey optimizasyona dayalı makine öğrenmesi algoritmalarıdır [34, 35, 36, 37]. Bu algoritma, veriye ilişkin herhangi bir birleşik dağılım fonksiyonu bilgisine ihtiyaç duymadığı içi dağılımdan bağımsız öğrenme algoritmalarıdır.

Destek Vektör Makineleri, örüntü tanıma ve sınıflandırma problemlerinin çözümü için Vapnik tarafından geliştirilmiştir. Destek Vektör Makinelerinin temelleri istatistiksel öğrenme teorisine diğer bir ifadeyle Vapnik-Chervonenkis (VC) teoresine dayanmaktadır. Şekil 2.13'da DVM'nin ağ yapısı gösterilmiştir.



Şekil 2.13 Destek Vektör Makinesi Yapısı [37]

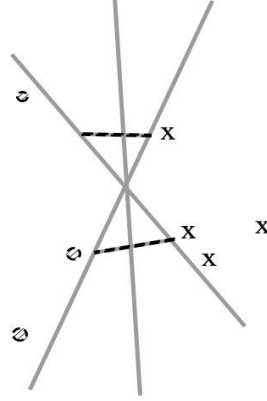
Şekil 2.14’de gösterilen ağ yapısı incelendiğinde,  $K(X_i, X_j)$  çekirdek fonksiyonları ve  $\alpha$  Lagrange çarpanlarını göstermektedir. Çekirdek fonksiyonları yardımı ile girdilerin iç çarpımları hesaplanmaktadır. Lagrange çarpanları ise ağırlıkları göstermektedir. DVM’lerde bir örneğe ait çıktı değeri, girdilerin iç çarpımları ile Lagrange çarpanlarının bağımsız kombinasyonlarının toplamına eşittir [37]. DVM’lerde amaç, sınıfları birbirinden ayıracak optimal ayırma hiper düzleminin elde edilmesidir. Başka bir ifadeyle, farklı sınıflara ilişkin destek vektörleri arasındaki uzaklığı maksimize etmektir. Destek Vektör Makineleri iki sınıflı veya çok sınıflı sınıflandırma probleminin çözümü için geliştirilmiş bir makine dili öğrenme algoritmasıdır. Bu çalışmada 2 sınıflı sınıflandırma algoritması üzerine odaklanılmıştır. Bu yüzden iki sınıflı Destek Vektör Makinesi’nin matematiksel yapısı anlatılmıştır.

Destek Vektör Makinesi veri setinin doğrusal olarak ayrılıp ayrılamama durumuna göre temel olarak ikiye ayrılmaktadır. Dolayısıyla çalışmanın bu bölümünde doğrusal ve doğrusal olmayan olmak üzere iki kategoride ele alınmıştır.

### Doğrusal Destek Vektör Makineleri

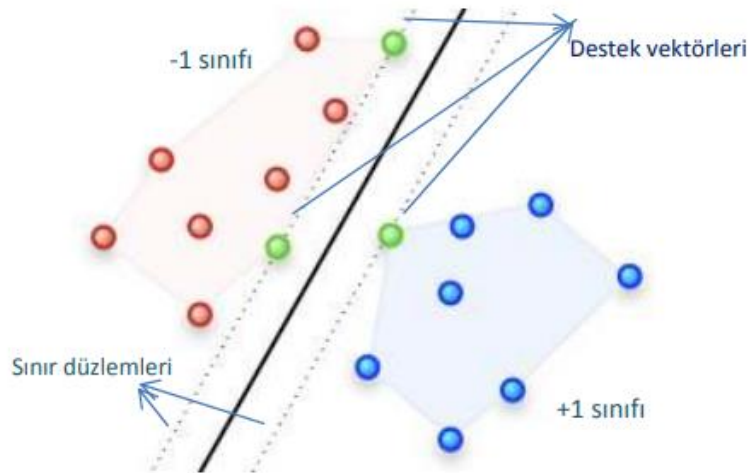
Her  $x_i$  örneği,  $p$  adet niteliğe sahip girdi,  $y_i \in \{-1,1\}$  örneklerinin ait olduğu sınıfı temsil eden çıktı ve  $x \in p$  yüksek boyutlu girdi vektörü olmak üzere;  $(x_i, y_i)$  ikililerinden oluşan  $n$  hacimli bir eğitim kümesi  $S$  verildiğinde, farklı sınıflara ait örnekleri birbirinden en iyi şekilde ayıracak,  $wx + b$  doğrusal hiper düzleminin bulunmasına yardımcı olacak denetimli öğrenme algoritmaları sınıfına ait makine

öğrenimi algoritmasıdır [37]. Burada  $w$ , hiper düzlemin normali aynı zamanda ağırlık vektörü ve  $b$  sayısı ise sabit bir katsayı olarak tanımlanmıştır.



Şekil 2.14 2 Sınıflı Veri Setini Ayıran Farklı Düzlemlere İlişkin Örnek [37].

Veri setini ayıran düzlemlere ilişkin geometrik gösterim Şekil 2.14'de gösterilmiştir. Şekilde görüldüğü gibi farklı sınıflara ait örnekleri birbirinden ayıran birçok doğrusal düzlem bulunabilir. Ancak DVM, farklı sınıflara ait destek vektörleri arasındaki uzaklığı maksimize eden ayırma hiper düzleminin bulunmasını amaçlar.



Şekil 2.15 Destek Vektör Makinesi için 2 sınıflı Problem Örneği [37]

2 sınıflı ve iki boyutlu bir sınıflandırma problemi için doğrusal DVM'nin geometrik gösterimi Şekil 2.15'de gösterilmiştir. Destek vektörleri, ayırma hiper düzlemine en yakın olan her iki sınıfa da ait örnekler olarak ifade edilmiş ve Şekil 2.16'da

gösterilmiştir. Söz konusu destek vektörleri, ait olduğu sınıfın sınırını belirler ve ayırma hiper düzlemine paralel bir düzlem üzerinde yer alır [42]. Destek vektörlerinin üzerinde bulunduğu ve kesikli çizgilerle gösterilmiş düzlemlere sınır düzlemleri denmektedir. Sınır düzlemlerinin tam ortasından geçen ve her iki düzleme de eşit uzaklıkta bulunan düzlem ise hiper düzlem olarak tanımlanmaktadır.

DVM veri setinin doğrusal ayrılma ve doğrusal olmayan ayrılma durumu olarak ikiye ayrılmaktadır.

Doğrusal Ayrılma Durumu; eğitim veri setinin doğrusal olarak ayrılabilme durumunda Destek Vektör Makineleri, en büyük sınıra sahip ayırma hiper düzlemini bulmaya çalışır. Söz konusu ayırma hiper düzleminin bulunabilmesi için veri setindeki tüm örneklerin,

$$f(X_1) = (w, X_1) + b \geq +1, \quad Y_1 = +1$$

$$f(X_1) = (w, X_1) + b \leq -1, \quad Y_1 = -1$$

eşitsizliklerini sağlaması gerekir [37]. Bu eşitsizlikler, tek bir eşitsizlikte birleştirilebilir.

$$\forall_i \text{ için } Y_i((w, x_i) + b) - 1 \geq 0$$

Bir hiper düzlem,  $w$  normali ve orijinden dik uzaklığı  $|b|/|w|$  olan,  $w \cdot x + b = 0$  düzlemdir.

Yukarıdaki eşitsizlikler göz önüne alındığında, sırasıyla,  $w$  normali ve orijinden dik uzaklığı  $|1 - b|/|w|$  olan  $H_1 = \langle w, x_i \rangle + b = +1$  düzlemi ile  $w$  normali ve orijinden dik uzaklığı  $|-1 - b|/|w|$  olan  $H_2 = \langle w, x_i \rangle + b = -1$  düzlemi paralel düzlemlerdir. Dolayısı ile  $H_1$  ve  $H_2$  sınır düzlemleri hiper düzleme eşit uzaklıkta yer almaktadır.  $H_1$  ve  $H_2$  sınır düzlemleri arasında herhangi bir eğitim örneği yer almamaktadır. Ancak, düzlemlerin üzerinde var olan eğitim örnekleri destek vektörleridir ve hiper düzleme en yakın olan eğitim örnekleridir. Ayırma hiper düzlemi, her iki sınıfın destek vektörleri arasındaki uzaklığı diğer bir ifade ile sınırı maksimum yapan ve sınırın ortasından geçen düzlemdir. Bir eğitim örneğinin hiper düzleme olan uzaklığı Eşitlik 2.19'da gösterilmiştir.

$$d(w, b; x) = \frac{|(w, x_i) + b|}{\|w\|} \quad (2.19)$$

Burada  $\|w\|$ , ağırlık vektörü olarak adlandırılan  $w$  normal düzleminin normudur. Dolayısı ile hiper düzleme en yakın örneklerin hiper düzleme olan ters uzaklığı sahip olduğu ağırlık vektörünün normuna eşit olmak zorundadır [40].

Bu teoremden yola çıkarak, eğitim örneklerini en iyi şekilde ayıran hiper düzlem,

$$\theta(w) = \frac{1}{2} \|w\|^2 \text{ eşitliğini en küçükleyen düzlemdir.}$$

Destek Vektör Makinelerinde optimal ayırma hiper düzleminin bulunması için kareli optimizasyon probleminin çözümünde Lagrange çarpanlarından faydalanılır. Lagrange çarpanları en küçükleme tipindeki problemi dual probleme dönüştürerek problemin daha kolay çözülmesine olanak sağlar. Problemin çözümünde kullanılan Lagrange fonksiyonu Eşitlik 2.20'de gösterilmiştir.

$$L_p(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i((w, x_i) + b)] + \sum_{i=1}^n \alpha_i \quad (2.20)$$

$\alpha_i$ 'nin çözümü ile optimal hiper düzlemin Eşitlik 2.21 ve 2.22'de verilen ağırlık vektörü  $w$  ve  $b$  sabit parametreleri hesaplanır.

$$w^* = \sum_{i=1}^n y_i \alpha_i x_i \quad (2.21)$$

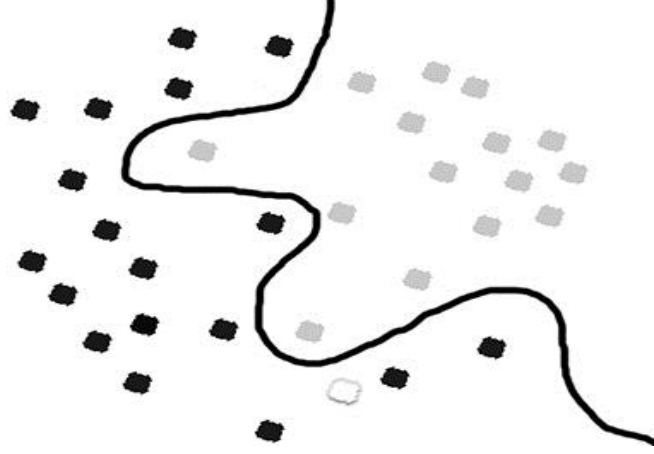
$$b^* = -\frac{1}{2} (w^*, x_s) \quad (2.22)$$

Sonuç olarak elde edilen hiper düzleme bağlı olarak sınıflayıcı Eşitlik 2.23'de verilmiştir.

$$f(x) = \text{sgn}((w^*, x_i) + b^*) = \text{sign}(\sum_{i=1}^n y_i \alpha_i (x_i, x_j)) \quad (2.23)$$

Doğrusal Olmayan Destek Vektör Makineleri; veri setinin doğrusal bir fonksiyon ile tam olarak veya belirli bir hata ile ayıramaması durumunda kullanılan algoritmalarıdır. Gerçek yaşam problemlerinde bir veri setinin hiper düzlem ile doğrusal olarak ayrılması genelde mümkün değildir. Dolayısı ile sınıflara ayırma işlemi, ayırma eğrisinin tahmin edilmesiyle mümkün olmaktadır. Ancak uygulamada eğrinin tahmin edilmesi oldukça zordur.

Veri setinin doğrusal olarak ayrılamama durumunun geometrik gösterimi Şekil 2.16'da verilmiştir. Bu durumda p-boyutlu girdi vektörü X'in p-boyutlu özellik vektörü  $\Phi$ 'ye dönüştürülmesi gerekmektedir [41].



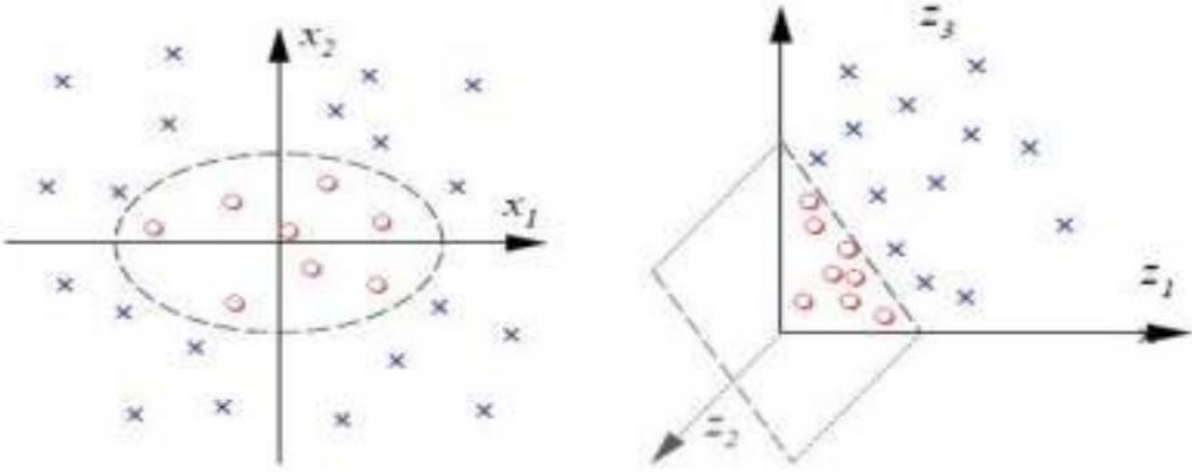
Şekil 2.16 2 Sınıflı Problem İçin Doğrusal Ayrılamama Durumu

P-boyutlu girdi vektörü X'in p-boyutlu özellik vektörü  $\Phi$ 'ye dönüştürülebilmesi için optimal ayırma düzleminin özellik uzayında tanımlanabilmesi gerekir. Bu amacı gerçekleştirmek için doğrusal olmayan haritalama yaklaşımından yararlanılır [42].

"Doğrusal olmayan haritalama", orijinal girdi uzayı X' in bir Hilbert uzayı olan daha yüksek boyutlu F özellik uzayına dönüştürülerek doğrusal ayrıştırımının gerçekleştirilmesi için kullanılan bir yaklaşımdır [43,44,45].

"Hilbert uzayı" pozitif skaler çarpıma sahip ve öğeleri fonksiyonlardan oluşan tam iç çarpım uzayları olarak ifade edilmektedir. Doğrusal olarak ayrılamayan iki boyutlu bir veri seti için doğrusal olmayan haritalama yaklaşımının geometrik olarak açıklaması Şekil 2.18'de verilmiştir. Doğrusal olmayan haritalama yaklaşımı ile iki boyutlu veri seti üç boyutlu özellik uzayına taşınarak veri setinin doğrusal ayrımı sağlanmaktadır.





Şekil 2.17 Doğrusal Olmayan Haritalama Yaklaşımı Örneği [37]

Şekil 2.17’de, iki boyutlu girdi vektörü  $(X_1, X_2)$  ve üç boyutlu özellik uzayı  $(z_1, z_2, z_3)$  olmak üzere, fonksiyonların özellik uzayı, Eşitlik 2.24’de verilmektedir.

$$\emptyset = R^2 \rightarrow R^3(X_1, X_2) \rightarrow (Z_1, Z_2, Z_3) := (X_1^2, \sqrt{2}X_1X_2, X_2^2) \quad (2.24)$$

Özellik uzayında haritalanmış girdi vektörlerinin iç çarpımları Eşitlik 2.25’deki gibi elde edilir.

$$\langle X, X' \rangle = (X_1^2 X_1'^2 + 2X_1 X_1' X_2 X_2' + X_2^2 X_2'^2)^2 \quad (2.25)$$

Böylece veri seti iki boyutlu uzaydan üç boyutlu uzaya taşınarak haritalama işlemi gerçekleştirilmiş olur. Sonuç olarak, doğrusal olmayan Destek Vektör Makineleri için özellik uzayında tanımlı ayırma hiper düzlemine bağlı olarak sınıflayıcı karar fonksiyonu Eşitlik 2.26’daki gibi ifade edilir.

$$f(x) = \text{sign}((w^*, X_i) + b^*) = \text{sign}\left(\sum_{i=1}^n y_i \alpha_i \left(\emptyset(x_i), \emptyset(x_j)\right)\right) \quad (2.26)$$

Özet olarak bu sınıflayıcının çalışma mantığı şu şekildedir: giriş vektörlerini önceden belirlenen doğrusal olmayan haritalama yöntemleriyle bazı yüksek boyutlu özellikleri Z uzayında haritalamaktır. Bu uzayda bazı özel özelliklerle birlikte bir doğrusal karar yüzeyi inşa edilir ve ağa yüksek bir genelleme yapma kabiliyeti kazandırılır. Bu genellemenin amacı bir dizi modelden yararlanılarak gelecek değerler için doğru bir yaklaşım yapabilmektir. Ayırıştırılan bu kategoriler mümkün olduğunca net bir şekilde ayrılmalıdır. Yeni gelen örnekler inşa edilen uzayda hangi tarafta yer alıyorsa o

kategoride yer alacağı tahmin edilir. Böylece oluşturulan destekçi vektör makinesi yeni gelen eğitim setlerini iki kategoriden bir tanesine atamaktadır. Bu sınıflayıcı kernel fonksiyonlarını kullanmaktadır. Bu fonksiyonlar girdilerini dolaylı yoldan yüksek boyutlu özellik uzayına haritalarlar ile temsil edilen her girdi, D özelliğine sahip olsun ve sadece = -1 ya da 1 sınıflarından birine ait olsun, bu durumda tüm girdileri şöyle gösterebiliriz:

Girilen parametrelere bakıldığında; yapılan bu çalışmanın sonucunda istenilen şey kişinin Parkinson hastası olup olmadığını anlamak olduğundan sınıf sayısı 2 olarak girilmiştir. Kernel yöntemi olarak Radial Basis Function (RBF) seçilmiştir [40]. Bu fonksiyon Eşitlik 2.27'de gösterilmiştir.

$$K(x,y) = e^{-\gamma|x-y|^2} \quad (2.27)$$

X vektörlerinin sınıflandırılması için kullanılan fonksiyonun matematiksel denklemi Eşitlik 2.28'de gösterilmiştir.

$$c = \sum_i a_i k(s_i, x) + b \quad (2.28)$$

Eşitlik 2.28'de yer alan  $s_i$  destek vektörlerini,  $a_i$  ağırlıkları, b ise bias değerlerini, k ise kernel fonksiyonunu göstermektedir. Doğrusal kernel durumlarında k nokta çarpımdır. Eğer  $c \geq 0$  ise X birinci grubun üyesi olarak sınıflandırılır. Aksi durumda ikinci grupta yer alır.

#### 2.4.2 k En Yakın Komşu (kEYK) (k Nearest Neighbor - kNN)

Parametrik olmayan kEYK metodu örüntü tanıma, sınıflandırma ve regresyon için kullanılmaktadır [38, 39]. Her iki durumda da giriş özellik uzayındaki k. en yakın eğitim örneklerinden oluşur. Çıkış ise kEYK'nın sınıflayıcı olarak mı yoksa regresyon için mi kullanılacağına bağlıdır.

kEYK sınıflandırmasında çıkış bir sınıf üyesidir. Bu sınıflandırma modelinde nesne komşuların oylarının çoğunluğuyla sınıflandırılır. Nesne k en yakın komşular arasında en yaygın olan sınıfa atanır. Eğer k=1 ise, nesne sadece o en yakın komşunun sınıfına atanır.

kEYK örnek tabanlı tembel bir sınıflayıcıdır. Denetimli öğrenme ile çalışır. Tüm makine dili öğrenme algoritmalarının en basitidir.

Algoritma kısmına bakıldığında, eğitim örnekleri, her biri bir sınıf etiketine sahip çok boyutlu bir özellik uzayındaki vektörlerdir. Algoritmanın eğitim fazı özellik vektörlerinin depolanması ve eğitim örneklerinin sınıflara etiketlendirilmesinden oluşur. Bir örnek komşuların çoğunluk oyuyla sınıflandırılır. Bu olay bir mesafe fonksiyonu ile ölçülen en yakın komşular arasında en yakın olan sınıfa atanır. Mesafe fonksiyonu için birden çok seçenek mevcuttur. Bunlardan bazıları şunlardır: Öklit, Manhattan, Minkowski. Bu üç mesafe fonksiyonunda da yalnızca sürekli değişkenler için geçerli olduğuna dikkat edilmelidir. Yapılan bu çalışmada Öklit mesafe fonksiyonu kullanılmıştır. Öklit mesafe fonksiyonunun matematiksel hesaplama formülü 2.29 ve 2.30'da gösterilmiştir.

$$\sqrt{\sum_{i=1}^k (x_i - x'_i)^2} \quad (2.29)$$

$$d(x, x') = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \dots + (x_n - x'_n)^2} \quad (2.30)$$

Eşitlik 2.29 ve 2.30'daki Öklit mesafe fonksiyonun haricinde Manhattan, Chebyshev ve Hamming gibi farklı mesafe fonksiyonları da kullanılabilir. Mesafe fonksiyonu içerisinde bulunan k parametresi kullanıcıya ve datanın kendisine bağlıdır. Bu değer çok büyük olmayacak şekilde seçilmesine özen gösterilir. Eğer binary sınıflandırma yapılacaksa en az 3 seçilir. Küçük k değerlerinde sonuç gürültüden daha çok etkilenmektedir. Bu değer büyüdükçe ise hesaplama süresi ve maliyet artmaktadır. Yine aynı şekilde küçük k değerlerinde bias az, varyans yüksek olur. Yüksek k değerlerinde ise bias yüksek, varyans düşük olmaktadır. Yapılan bu çalışmada k değeri 7 olarak alınmıştır. Eğer etiketlenmemiş bir vektör varsa sorgu noktasına en yakın k eğitim örnekleri arasında en sık kullanılan etiketin atanmasıyla sınıflandırılır. Yapılan bu çalışmada etiketlenmemiş bir vektör olmamasına dikkat edilmiştir. Yapılan bu çalışmada k değerleri birden yirmiye kadar denenmiştir.

Bu sınıflayıcı modelinde en yakın k adet komşunun her birinin ağırlıkları  $\frac{1}{k}$  olarak hesaplanır, geriye kalanlar ise 0 olarak alınmaktadır.

Daha usulen anlatmak gerekirse,  $k$  pozitif bir tam sayı,  $x$  görünmeyen bir gözlem ve  $d$  bir benzerlik metriği verildiğinde bu sınıflayıcı türü aşağıdaki iki adımı gerçekleştirir:

- $X$  ile her bir eğitim gözlemi arasındaki tüm veri setlerinin  $d$  mesafeleri hesaplanır. Eğitim seti içerisinde  $A$  seti içerisindeki  $x$ 'e en yakın  $k$  noktası aranılır. Bağlantı noktalarındaki sorunları aşmak için  $k$  genellikle tek sayı seçilir.
- Daha sonra her bir sınıf için koşullu olasılık tahmini yapılır. Yani verilen sınıf etiketlerinin  $A$  seti içerisindeki noktaların oranı hesaplanır.

$$P(Y = j | X = x) = \frac{1}{k} \sum_{i \in A} I(y^{(i)} = j) \quad (2.31)$$

Eşitlik 2.31 ile koşullu olasılık hesaplaması yapılmıştır.  $I(x)$  işaret fonksiyonudur.  $X$  argümanı doğru olduğu durumlarda 1, yanlış olduğu durumlarda 0 döndürür.

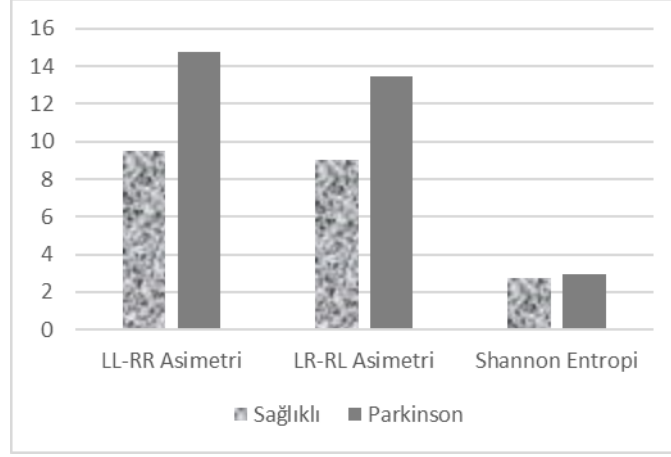
### 3. SONUÇLAR

#### 3.1 İstatistiksel Analiz Sonuçları

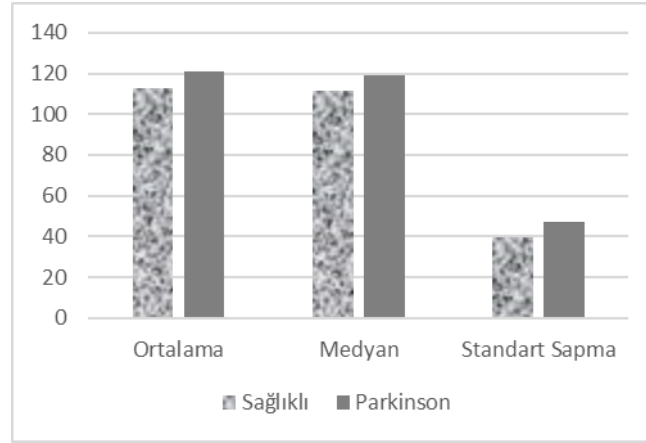
Bireylere ait toplamda 551 adet kayıt bulunmaktadır. Kayıtlardaki örnek sayıları sabit olmamakla birlikte bazı kayıtlar çok az örnek içermektedir. Hesaplanan özelliklerin örnek sayısına bağımlılığını ortadan kaldırmak için ilk olarak örnek sayısı 1000'in altında olan kayıtlar çalışmaya dahil edilmemiş (bu durumda 515 kayıt üzerinde analizler gerçekleştirilmiştir), ikinci olarak da çok uzun örnekler içeren kayıtlar 1000 civarı örnek içerecek şekilde bölünerek kullanılmıştır (böylece 646 kayıt elde edilmiştir). Örnek azlığı nedeniyle çıkarılan kayıtlar dikkate alındığında katılımcı sayısı; 147 tanesi Parkinson hastası, 49 tanesi de sağlıklı birey olmak üzere 196 olmuştur. Histogram alanları toplam 50 bölgede kontrol edildiği için ayırt edici olarak bulunan bölgeler her birinde farklı olabileceğinden dolayı tüm özelliklerden bağımsız olarak farklı bir tabloda gösterilmiştir. İstatistiksel analiz sonuçları ANOVA'ya göre yapılarak elde edilmiş ve aşağıdaki tablolar aracılığı ile sunulmuştur. Aşağıdaki tablolarda ayırt edici olarak bulunan sonuçlar ( $P \leq 0,05$ ) koyu olarak işaretlenmiştir. 0,000 olarak belirtilen  $P$  değerleri,  $P \leq 0,0005$ 'i ifade etmektedir. \* ile işaretlenmiş olan değerler eşit varyans dağılımına sahip olmayanları göstermektedir.

Çizelge 3.1 Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre basılı tutma sürelerinin ANOVA sonuçları.

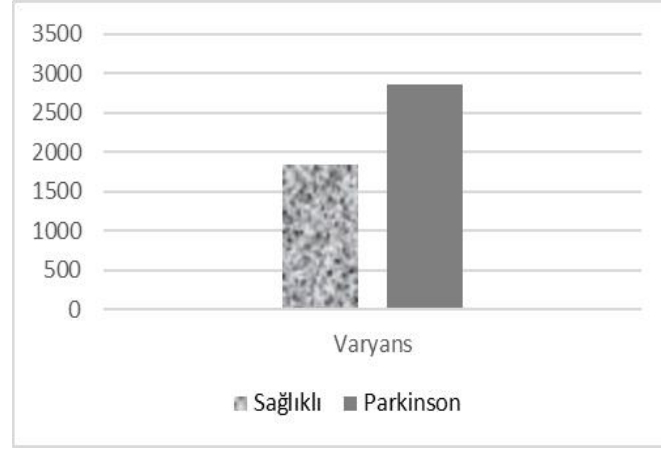
Özellik	Grup	Ortalama	Std. Sapma	Min.	Maks.	T	Sig (P)	
LL-RR	Normal	9,53	9,341	0,00	50	-3,999	<b>0,000</b>	
	Parkinson	14,80	14,80	0,00	135			
LR-RL	Normal	8,9994	10,2931	0,05	58,37	-4,251	<b>0,000</b>	
	Parkinson	13,4612	11,1747	0,00	58,88			
Basıklık	Normal	1,0101	2,5036	2,1958	4,0907	-0,915	0,361	
	Parkinson	2,8197	2,5709	1,7656	4,4702			
Çarpıklık	Normal	0,6683	1,1309	-0,6599	7,3637	-0,597	0,550	
	Parkinson	0,8565	3,6790	-50,012	49,5702			
İlk Fark Durum Uzayı	I.	Normal	0,2240	0,2260	0,0000	0,9993	-1,171	0,242
		Parkinson	0,2479	0,2102	0,0000	0,9968		
	II.	Normal	0,2674	0,0856	0,0003	0,3768	-1,278*	0,203*
		Parkinson	0,2775	0,0698	0,0016	0,3962		
	IV.	Normal	0,2419	0,2221	0,0000	0,9931	2,174*	<b>0,031*</b>
		Parkinson	0,1973	0,1868	0,0000	0,9039		
	III.	Normal	0,2667	0,0851	0,0003	0,3557	-1,362	0,175
		Parkinson	0,2773	0,0695	0,0016	0,4012		
Durum Uzayı	I.	Normal	0,0177	0,0428	0,0000	0,3484	0,720	0,472
		Parkinson	0,0156	0,0491	0,0000	0,6261		
	II.	Normal	0,0530	0,0597	0,0001	0,2416	1,163	0,245
		Parkinson	0,0426	0,0573	0,0001	0,2360		
	IV.	Normal	0,8761	0,1556	0,1992	0,9999	-1,094	0,274
		Parkinson	0,8992	0,1519	0,0562	0,9999		
	III.	Normal	0,0531	0,0597	0,0001	0,2416	1,172	0,242
		Parkinson	0,0426	0,0571	0,0001	0,2369		
Shannon Entopi	Normal	2,7660	0,6788	0,7693	3,8213	-3,359	<b>0,001</b>	
	Parkinson	2,9679	0,6149	0,9273	4,4727			
Örnek Entropi	Normal	1,7516	0,5562	0,0000	2,8507	-0,526	0,599	
	Parkinson	131,70	2919,1	0,0000	6,5535			
Yaklaşık Entropi	Normal	0,1648	0,1573	0,0052	0,9307	-3,639	<b>0,000</b>	
	Parkinson	0,2288	0,1910	-1,4142	0,9953			
Ortak Bilgi	Normal	4,0071	1,7690	2,0000	10,0000	-0,189	0,850	
	Parkinson	4,0395	1,7984	2,0000	18,0000			
Yerleştirme Boyutu	Normal	5,0071	1,1411	3,0000	10,0000	0,393	0,695	
	Parkinson	4,9644	1,1392	1,0000	10,0000			
Ortalama	Normal	112,6187	24,9470	65,5683	175,382	-3,272*	<b>0,001*</b>	
	Parkinson	121,0199	32,9659	54,1718	234,332			
Medyan	Normal	111,4250	25,2402	62,5000	171,900	-2,852*	<b>0,005*</b>	
	Parkinson	118,8506	33,5725	58,6000	234,400			
Standart Sapma	Normal	39,4837	16,8104	11,9109	96,320	-3,294	<b>0,001</b>	
	Parkinson	47,0262	25,6020	12,3252	329,287			
Varyans	Normal	1839,534	1436,757	141,869	9277,7	-0,754	<b>0,000</b>	
	Parkinson	2865,630	5601,536	151,910	10843			



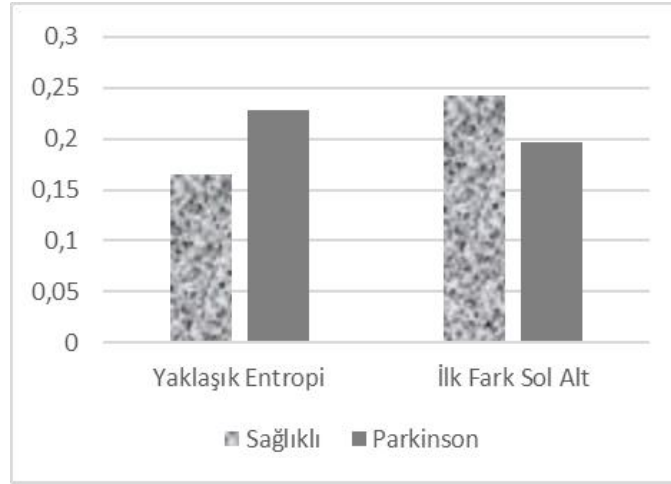
Şekil 3.1 646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan LL-RR Asimetri, LR-RL Asimetri ve Shannon Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



Şekil 3.2 646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Ortalama, Medyan ve Standart Sapma özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



Şekil 3.3 646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Varyans özelliğinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



Şekil 3.4 646 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Yaklaşık Entropi ve İlk Fark IV. özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.

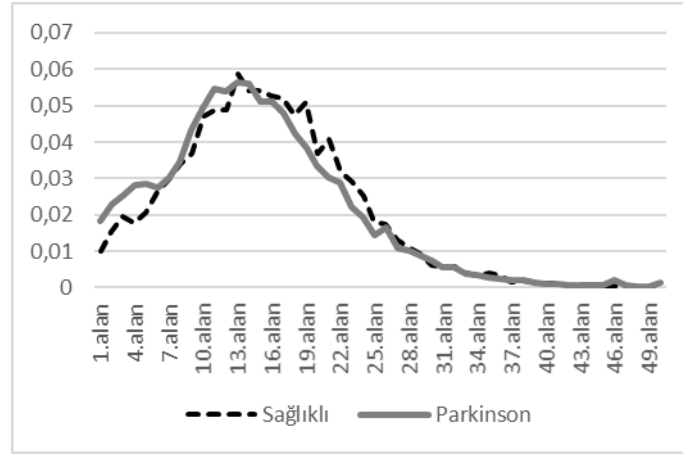
Çizelge 3.2 Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre basılı tutma sürelerinin Histogram alanlarının ANOVA sonuçları.

Özellik	Grup	Ortalama	Std. Sapma	Min.	Maks.	t	Sig (P)	
LL-RR	Normal	8,9169	8,3399	0,2696	46,070	-5,475*	<b>0,000*</b>	
	Parkinson	14,8317	14,9971	0,0379	135,211			
LR-RL	Normal	8,4780	10,0731	0,0511	58,374	-4,378*	<b>0,000*</b>	
	Parkinson	13,2329	10,9085	0,0075	58,880			
Basıklık	Normal	11,2096	10,4974	2,1958	24,907	-0,914	0,361	
	Parkinson	33,6394	26,4355	2,1024	4470,2			
Çarpıklık	Normal	0,7066	1,2209	-0,5518	7,3637	-0,405	0,503	
	Parkinson	0,8637	4,0994	-50,012	49,570			
İlk Fark Durum Uzayı (1 <sup>st</sup> Dif.)	I.	Normal	0,2240	0,2385	0,0000	0,9993	-1525	0,128
		Parkinson	0,2599	0,2180	0,0000	0,9968		
	II.	Normal	0,2623	0,0914	0,0003	0,3656	-1,475*	0,142*
		Parkinson	0,2759	0,0718	0,0016	0,3962		
	IV.	Normal	0,2520	0,2317	0,0000	0,9931	2,707*	<b>0,008*</b>
		Parkinson	0,1884	0,1838	0,0000	0,9039		
III.	Normal	0,2616	0,0908	0,0003	0,3557	-1,521*	0,130*	
	Parkinson	0,2757	0,0717	0,0016	0,4012			
Durum Uzayı	I.	Normal	0,0177	0,0428	0,0000	0,3484	0,420	0,675
		Parkinson	0,0156	0,0491	0,0000	0,6261		
	II.	Normal	0,0531	0,0596	0,0001	0,2416	1,706	0,089
		Parkinson	0,0426	0,0573	0,0001	0,2360		
	IV.	Normal	0,8761	0,1555	0,1992	0,9999	-1,431	0,153
		Parkinson	0,8992	0,1519	0,0562	0,9999		
III.	Normal	0,0531	0,0597	0,0001	0,2416	1,728	0,085	
	Parkinson	0,0426	0,0571	0,0001	0,2369			
Shannon Entropi	Normal	2,8148	0,6644	0,7693	3,8213	-3,057	<b>0,002</b>	
	Parkinson	3,0184	0,6194	0,9273	4,4727			
Örnek Entropi	Normal	1,8202	0,0449	0,0187	2,8507	1,917	0,056	
	Parkinson	1,7394	0,0185	0,0001	2,3912			
Yaklaşık Entropi	Normal	0,2458	0,1718	0,0005	0,8792	-4,132*	<b>0,000*</b>	
	Parkinson	0,2297	0,1713	0,0005	0,9307			
Ortak Bilgi	Normal	4,1913	1,8156	2,0000	10,0000	-0,173	0,863	
	Parkinson	4,2250	1,8541	2,0000	18,0000			
Yerleştirme Boyutu	Normal	5,2695	1,0622	4,0000	10,0000	-0,028	0,978	
	Parkinson	5,2725	0,9903	4,0000	10,0000			
Ortalama	Normal	111,5593	22,5656	66,4427	167,744	3,255*	<b>0,001*</b>	
	Parkinson	120,2779	33,1573	54,1718	231,872			
Medyan	Normal	110,4591	23,1501	62,5000	168,900	-2,869*	<b>0,004*</b>	
	Parkinson	118,3193	33,7375	58,6000	226,600			
Standart Sapma	Normal	40,7027	16,9458	14,5882	96,3208	-2,974	<b>0,003</b>	
	Parkinson	48,6825	27,2906	13,0358	329,287			
Varyans	Normal	1941,374	1489,57	212,815	9277,7	3,441*	<b>0,001*</b>	
	Parkinson	3112,905	6216,40	169,933	10843			

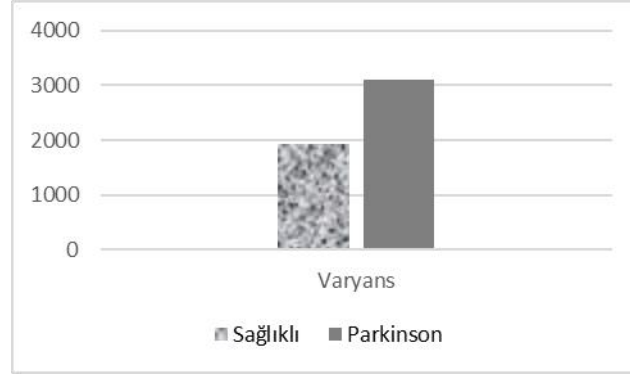


Çizelge 3.3 Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre basılı tutma sürelerinin ANOVA sonuçları.

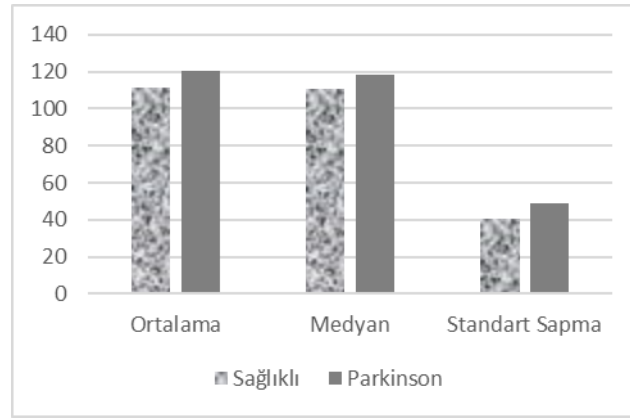
Özellik	Grup	Ortalama	Std. Sapma	Min.	Max.	T	Sig (P)
1. Alan	Normal	0,0174	0,0193	0,0000	0,1007	-3,033*	<b>0,003*</b>
	Parkinson	0,0186	0,0194	0,0000	0,1056		
4. Alan	Normal	0,0370	0,0342	0,0000	0,2424	-2,474*	<b>0,014*</b>
	Parkinson	0,0311	0,0232	0,0000	0,1756		
19. Alan	Normal	0,0284	0,0135	0,0055	0,0884	2,628*	<b>0,009*</b>
	Parkinson	0,0312	0,0134	0,0000	0,0701		
21. Alan	Normal	0,0256	0,0128	0,0044	0,0688	2,718*	<b>0,007*</b>
	Parkinson	0,0280	0,0123	0,0000	0,0700		
23. Alan	Normal	0,0219	0,0190	0,0000	0,0621	2,185	<b>0,029</b>
	Parkinson	0,0240	0,0133	0,0000	0,0806		



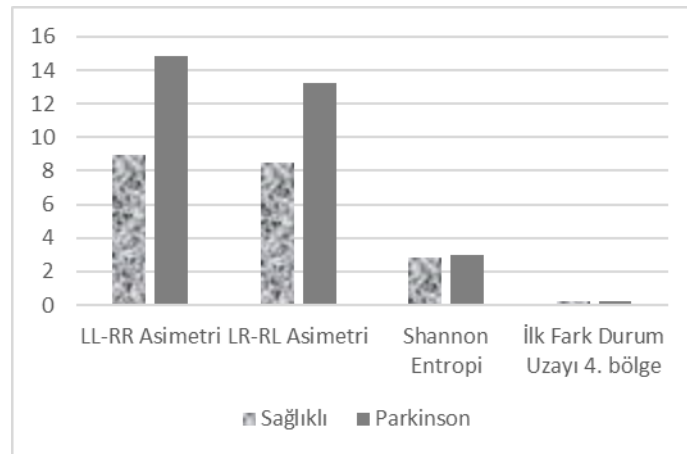
Şekil 3.5 646 kayda göre basılı tutma sürelerinin Histogram alanları



Şekil 3.6 515 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Varyans özelliğinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



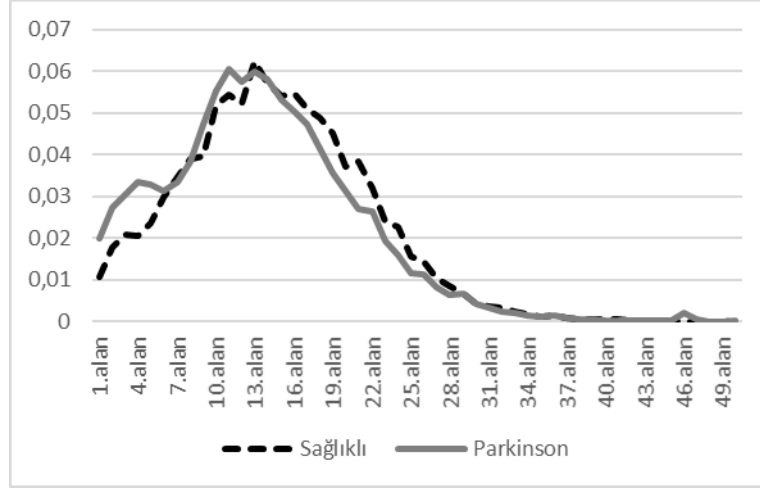
Şekil 3.7 515 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan Ortalama, Medyan ve Standart Sapma özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



Şekil 3.8 515 kayda göre basılı tutma sürelerinde ayırt edici olarak bulunan LL-RR Asimetri, LR-RL Asimetri, Shannon Entropi ve İlk Fark 4. Bölge için sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.

Çizelge 3.4 Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre basılı tutma sürelerinin Histogram alanlarının ANOVA sonuçları.

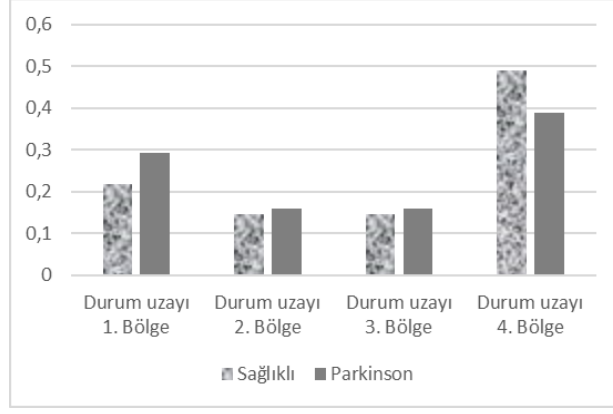
Özellik	Grup	Ortalama	Std. Sapma	Min.	Max.	T	Sig (P)
1. Alan	Normal	0,0088	0,0156706	0,0000	0,0811	-3,431*	<b>0,001*</b>
	Parkinson	0,0181	0,0352899	0,0000	0,2545		



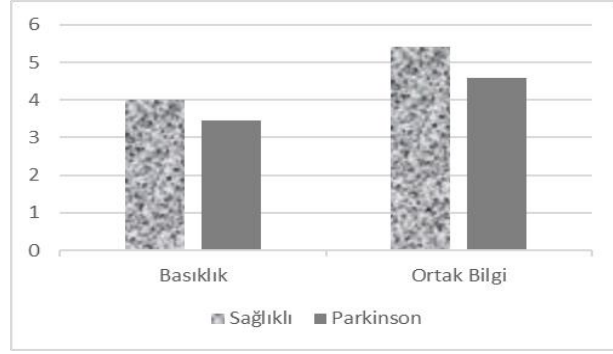
Şekil 3.9 515 kayda göre basılı tutma sürelerinin Histogram alanları.

Çizelge 3.5 Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre gecikme sürelerinin ANOVA sonuçları.

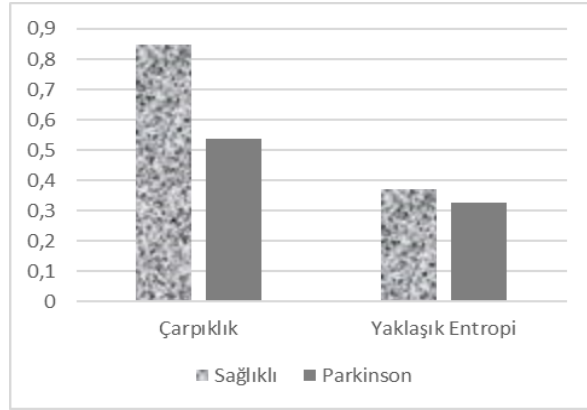
Özellik	Grup	Ortalama	Std. Sapma	Min.	Maks.	T	Sig (P)	
LL-RR	Normal	24,7120	26,1543	0,0500	222,64	-0,786*	0,433*	
	Parkinson	26,6245	22,9224	0,0600	142,35			
LR-RL	Normal	25,2665	24,2224	0,5100	137,51	-1,000	0,318	
	Parkinson	28,2681	33,1592	0,0400	307,97			
Basıklık	Normal	4,0050	1,9040	1,9013	8,9629	3,009	<b>0,003</b>	
	Parkinson	3,4430	1,9699	1,5658	17,0960			
Çarpıklık	Normal	0,8473	0,6734	-0,6124	2,1847	4,832	<b>0,000</b>	
	Parkinson	0,5360	0,6981	-2,8743	3,1779			
İlk Fark Durum Uzaı	I.	Normal	0,1863	0,1011	0,0421	0,5400	0,598	0,550
		Parkinson	0,1802	0,1083	0,0126	0,6003		
	II.	Normal	0,3077	0,0304	0,2147	0,4286	0,323	0,747
		Parkinson	0,3067	0,0306	0,1714	0,4278		
	IV.	Normal	0,2019	0,1049	0,0305	0,4687	0,196	0,435
		Parkinson	0,2101	0,1115	0,0063	0,6335		
III.	Normal	0,3041	0,0280	0,2148	0,3607	0,111	0,655	
	Parkinson	0,3030	0,0272	0,1714	0,3929			
Durum Uzaı	I.	Normal	0,2185	0,2184	0,0049	0,9853	3,476*	<b>0,001*</b>
		Parkinson	0,2930	0,2456	0,0000	0,9800		
	II.	Normal	0,1462	0,0610	0,0074	0,2685	-2,254*	<b>0,025*</b>
		Parkinson	0,1590	0,0549	0,0099	0,2892		
	IV.	Normal	0,4890	0,2931	0,0000	0,9036	3,663*	<b>0,000*</b>
		Parkinson	0,3888	0,2607	0,0000	0,9534		
III.	Normal	0,1462	0,0611	0,0074	0,2724	-2,237*	<b>0,026*</b>	
	Parkinson	0,1590	0,054	0,0099	0,2892			
Shannon Entopi	Normal	4,3626	0,3772	3,5768	5,1513	-0,010	0,992	
	Parkinson	4,3395	0,3639	3,3670	5,1351			
Örnek Entopi	Normal	1,6687	0,4879	0,0000	3,2095	-0,526	0,599	
	Parkinson	1,3125	2,9133	0,0000	6,5535			
Yaklaşık Entopi	Normal	0,3712	0,2193	0,0092	0,8187	2,294	<b>0,022</b>	
	Parkinson	0,3271	0,1963	-0,0136	0,9988			
Ortak Bilgi	Normal	5,4143	2,7874	2,0000	15,0000	3,924	<b>0,000</b>	
	Parkinson	4,5771	2,0564	2,0000	17,0000			
Yerleşirme Boyutu	Normal	4,5429	0,7715	2,0000	6,0000	-0,196	0,845	
	Parkinson	4,5573	0,7718	2,000	8,0000			
Ortalama	Normal	280,182	86,4813	152,179	575,266	-0,745	0,457	
	Parkinson	285,499	94,6204	145,436	656,445			
Medyan	Normal	259,744	91,7960	140,600	601,600	-1,246	0,213	
	Parkinson	269,597	100,480	117,900	671,900			
Standart Sapma	Normal	119,296	31,6260	64,9475	199,951	1,553	0,121	
	Parkinson	115,722	31,7089	73,7765	237,096			
Varyans	Normal	15223,1	7942,91	4218,2	39981	1,437	0,151	
	Parkinson	14394,7	8298,13	5443,0	56215			



Şekil 3.10 646 kayda göre gecikme sürelerinde ayırt edici olarak bulunan durum uzayı 1., 2., 3. ve 4. bölge nokta sayılarının Parkinson hastası ve sağlıklı bireyler için ortalama değerleri.



Şekil 3.11 646 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Basıklık ve Ortak Bilgi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



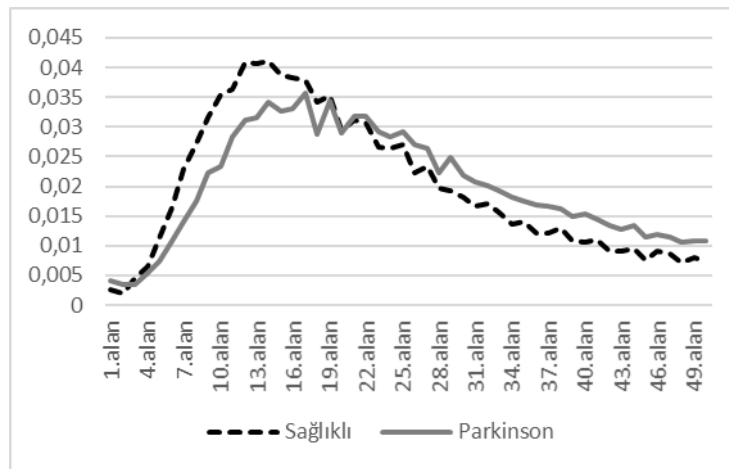
Şekil 3.12 646 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Çarpıklık ve Yaklaşık Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.

Çizelge 3.6 Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre gecikme sürelerinin Histogram alanlarının ANOVA sonuçlarını göstermektedir.

Özellik	Grup	Ortalama	Std. Sapma	Min.	Max.	t	Sig.
2. Alan	Normal	0,0020	0,0028	0,0000	0,0159	2,777	<b>0,006</b>
	Parkinson	0,0035	0,0076	0,0000	0,0690		
5. Alan	Normal	0,0114	0,0169	0,0000	0,1006	2,613	<b>0,010</b>
	Parkinson	0,0073	0,0129	0,0000	0,0972		
6. Alan	Normal	0,0165	0,0210	0,0000	0,0989	2,945	<b>0,040</b>
	Parkinson	0,0109	0,0159	0,0000	0,1040		
7. Alan	Normal	0,0230	0,0270	0,0000	0,1421	3,655	<b>0,000</b>
	Parkinson	0,0140	0,2033	0,0000	0,1159		
8. Alan	Normal	0,0273	0,0273	0,0000	0,1061	3,846	<b>0,000</b>
	Parkinson	0,0176	0,0226	0,0000	0,1066		
9. Alan	Normal	0,0317	0,0293	0,0000	0,0976	3,473	<b>0,001</b>
	Parkinson	0,0222	0,0210	0,0000	0,1470		
10. Alan	Normal	0,0354	0,0288	0,0000	0,1125	3,472	<b>0,000</b>
	Parkinson	0,0234	0,0255	0,0000	0,1618		
11. Alan	Normal	0,0363	0,0266	0,0000	0,1057	3,145	<b>0,002</b>
	Parkinson	0,0283	0,0265	0,0000	0,1206		
12. Alan	Normal	0,0408	0,0285	0,0000	0,1379	3,613	<b>0,000</b>
	Parkinson	0,0311	0,0261	0,0000	0,1103		
13. Alan	Normal	0,0406	0,0298	0,0000	0,1216	3,239	<b>0,001</b>
	Parkinson	0,0316	0,0258	0,0000	0,1270		
14. Alan	Normal	0,0410	0,0247	0,0000	0,1578	2,990	<b>0,003</b>
	Parkinson	0,3411	0,0238	0,0000	0,1051		
15. Alan	Normal	0,0387	0,0241	0,0000	0,1570	2,912	<b>0,004</b>
	Parkinson	0,0326	0,0214	0,0000	0,1111		
16. Alan	Normal	0,0384	0,0191	0,0000	0,1060	2,677	<b>0,008</b>
	Parkinson	0,0330	0,0214	0,0000	0,1106		
18. Alan	Normal	0,0341	0,0145	0,0000	0,0756	1,192	0,234
	Parkinson	0,0289	0,0175	0,0000	0,1290		
26. Alan	Normal	0,0222	0,0148	0,0000	0,0812	3,406	<b>0,001</b>
	Parkinson	0,0270	0,0150	0,0000	0,0862		
27. Alan	Normal	0,0234	0,0153	0,0000	0,0786	2,192	<b>0,029</b>
	Parkinson	0,0265	0,0141	0,0000	0,0857		
28. Alan	Normal	0,0198	0,0134	0,0000	0,0590	2,101	<b>0,036</b>
	Parkinson	0,0223	0,0129	0,0000	0,0643		
29. Alan	Normal	0,0193	0,0140	0,0000	0,0635	4,183	<b>0,000</b>
	Parkinson	0,0249	0,0139	0,0000	0,0736		
30. Alan	Normal	0,0181	0,0130	0,0029	0,0637	2,913	<b>0,004</b>
	Parkinson	0,0219	0,0138	0,0000	0,1014		
31. Alan	Normal	0,0166	0,0129	0,0000	0,0653	3,515	<b>0,000</b>
	Parkinson	0,0207	0,0122	0,0000	0,0641		
32. Alan	Normal	0,0171	0,0130	0,0000	0,0561	2,610	<b>0,009</b>
	Parkinson	0,0202	0,0124	0,0000	0,0598		
33. Alan	Normal	0,0156	0,0119	0,0000	0,0565	2,722	<b>0,007</b>
	Parkinson	0,0192	0,0139	0,0000	0,0738		
34. Alan	Normal	0,0137	0,0108	0,0000	0,0472	3,817	<b>0,000</b>
	Parkinson	0,0181	0,0123	0,0000	0,0797		
35. Alan	Normal	0,0141	0,0120	0,0000	0,0527	2,832	<b>0,005</b>
	Parkinson	0,0175	0,0129	0,0000	0,0870		

Çizelge 3.6 devam ediyor.

36. Alan	Normal	0,0122	0,0104	0,0000	0,0495	4,173	<b>0,000</b>
	Parkinson	0,0169	0,0120	0,0000	0,0646		
37. Alan	Normal	0,0121	0,0094	0,0000	0,0412	4,675	<b>0,000</b>
	Parkinson	0,0166	0,0126	0,0000	0,0725		
38. Alan	Normal	0,0129	0,0112	0,0000	0,0477	2,639	<b>0,009</b>
	Parkinson	0,0162	0,0134	0,0000	0,0888		
39. Alan	Normal	0,0108	0,0104	0,0000	0,0635	3,900	<b>0,000</b>
	Parkinson	0,0149	0,0124	0,0000	0,00597		
40. Alan	Normal	0,0107	0,0109	0,0000	0,0590	4,155	<b>0,000</b>
	Parkinson	0,0153	0,0141	0,0000	0,0961		
41. Alan	Normal	0,0110	0,0102	0,0000	0,0472	3,358	<b>0,001</b>
	Parkinson	0,0144	0,0125	0,0000	0,0667		
42. Alan	Normal	0,0100	0,0090	0,0000	0,0000	4,408	<b>0,000</b>
	Parkinson	0,0100	0,0130	0,0000	0,0000		
43. Alan	Normal	0,0091	0,0090	0,0000	0,0400	3,939	<b>0,000</b>
	Parkinson	0,0128	0,0124	0,0000	0,0800		
44. Alan	Normal	0,0100	0,0100	0,0000	0,0000	3,761	<b>0,000</b>
	Parkinson	0,0100	0,0130	0,0000	0,0000		
45. Alan	Normal	0,0100	0,0080	0,0000	0,0000	4,769	<b>0,000</b>
	Parkinson	0,0100	0,0110	0,0000	0,0000		
46. Alan	Normal	0,0100	0,0090	0,0000	0,0000	3,172	<b>0,002</b>
	Parkinson	0,0100	0,0120	0,0000	0,0000		
47. Alan	Normal	0,0087	0,0109	0,0000	0,0800	2,539	<b>0,012</b>
	Parkinson	0,0114	0,0127	0,0000	0,0800		
48. Alan	Normal	0,0100	0,0080	0,0000	0,0000	4,116	<b>0,000</b>
	Parkinson	0,0100	0,0110	0,0000	0,0000		
49. Alan	Normal	0,0100	0,0090	0,0000	0,0000	3,164	<b>0,002</b>
	Parkinson	0,0100	0,0110	0,0000	0,0000		
50. Alan	Normal	0,0100	0,0080	0,0000	0,0000	4,364	<b>0,000</b>
	Parkinson	0,0100	0,0110	0,0000	0,0000		

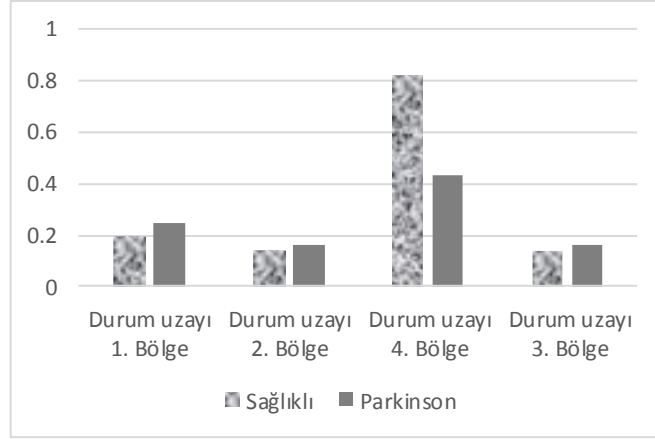


Şekil 3.13 646 kayda göre gecikme sürelerinin Histogram alanları.

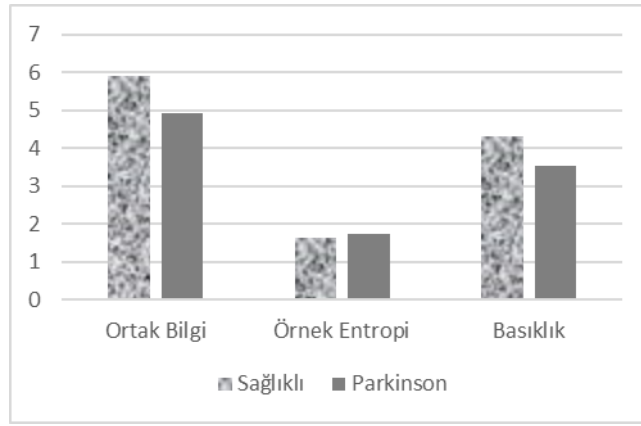
Çizelge 3.7 Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre gecikme sürelerinin ANOVA sonuçlarını göstermektedir.

Özellik	Grup	Ortalama	Std. Sapma	Min.	Maks.	T	Sig (P)	
LL-RR	Normal	21,5499	18,8417	0,0527	99,3914	-1,375	0,170	
	Parkinson	24,3772	19,5923	0,0631	96,5056			
LR-RL	Normal	20,2777	20,7897	0,5118	137,512	-1,174	0,241	
	Parkinson	22,9857	22,0798	0,0362	144,001			
Basıklık	Normal	4,3020	1,9584	1,9462	8,9629	3,704*	<b>0,000*</b>	
	Parkinson	3,5424	1,8659	1,6972	16,1651			
Çarpıklık	Normal	0,9389	0,6724	-0,6125	2,1847	4,415	<b>0,000</b>	
	Parkinson	0,6336	0,6480	-1,0695	2,7918			
İlk Fark Durum Uzaı	I.	Normal	0,1880	0,1038	0,0421	0,5400	1,113	0,266
		Parkinson	0,1758	0,1038	0,0197	0,6003		
	II.	Normal	0,3043	0,0302	0,2147	0,4286	-0,802	0,423
		Parkinson	0,3068	0,0282	0,1955	0,4278		
	IV.	Normal	0,2077	0,1097	0,0305	0,4687	-0,645	0,519
		Parkinson	0,2150	0,1056	0,0096	0,5867		
	III.	Normal	0,2998	0,0263	0,2148	0,3477	-0,927	0,355
		Parkinson	0,3023	0,0235	0,1946	0,3440		
Durum Uzaı	I.	Normal	0,1893	0,2004	0,0049	0,7884	-2,514	<b>0,012</b>
		Parkinson	0,2453	0,2131	0,0016	0,9800		
	II.	Normal	0,1374	0,0594	0,0457	0,2601	-3,869*	<b>0,000*</b>
		Parkinson	0,1612	0,0536	0,0099	0,2653		
	IV.	Normal	0,5358	0,2862	0,0216	0,9036	3,527*	<b>0,001*</b>
		Parkinson	0,4321	0,2458	0,0003	0,9534		
	III.	Normal	0,1374	0,0594	0,0457	0,2601	-3,876*	<b>0,000*</b>
		Parkinson	0,1612	0,0536	0,0099	0,2653		
Shannon Entopi	Normal	4,3413	0,3817	3,5768	5,1513	-0,027	0,978	
	Parkinson	4,3423	0,3495	3,5555	5,1351			
Örnek Entopi	Normal	1,6342	0,3472	0,3346	2,3631	-3,085	<b>0,002</b>	
	Parkinson	1,7385	0,3113	0,5167	2,6050			
Yaklaşık Entopi	Normal	0,4017	0,2106	0,0483	0,8187	1,983*	<b>0,049*</b>	
	Parkinson	0,3586	0,1859	0,0191	0,9988			
Ortak Bilgi	Normal	5,9217	2,7851	2,0000	15,0000	3,511*	<b>0,001*</b>	
	Parkinson	4,9375	2,1094	2,0000	17,0000			
Yerleştirme Boyutu	Normal	4,8000	0,5164	3,0000	6,0000	-0,557	0,578	
	Parkinson	4,8300	0,5067	3,0000	8,0000			
Ortalama	Normal	280,1826	86,4813	152,179	575,266	-0,541	0,589	
	Parkinson	285,4994	94,6204	145,436	656,445			
Medyan	Normal	259,7443	91,7959	140,600	601,600	-0,944	0,346	
	Parkinson	269,5970	100,489	117,9000	671,900			
Standart Sapma	Normal	119,2964	31,6259	64,9475	199,951	1,066	0,287	
	Parkinson	115,7228	31,7088	73,7765	237,096			
Varyans	Normal	152,2314	7942,91	4218,2	39981	0,952	0,341	
	Parkinson	143,9471	8298,13	5443,0	56215			

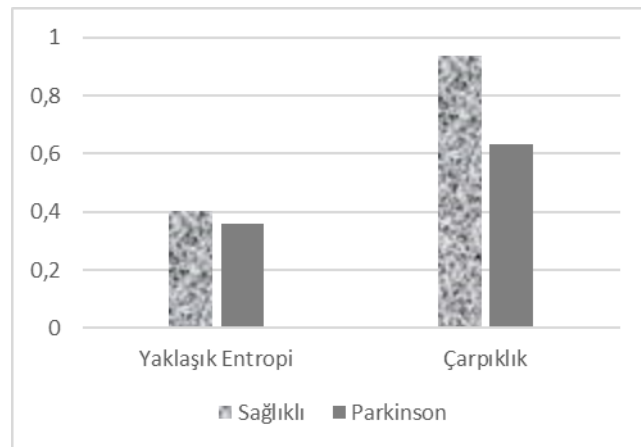




Şekil 3.14 515 kayda göre gecikme sürelerinde ayırt edici olarak bulunan durum uzayı 1., 2., 3. ve 4. bölge nokta sayılarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



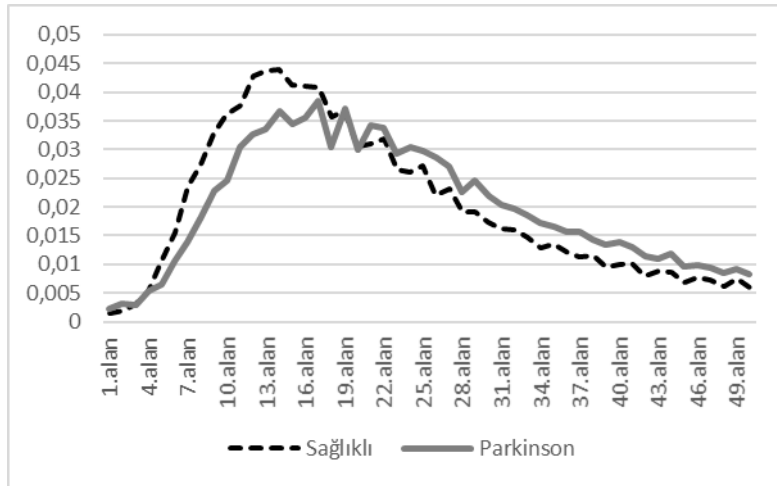
Şekil 3.15 515 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Basıklık, Örnek Entropi ve Ortak Bilgi özelliklerinin sağlıklı ve Parkinson hasta bireyler için ortalama değerleri.



Şekil 3.16 515 kayda göre gecikme sürelerinde ayırt edici olarak bulunan Yaklaşık Entropi ve Çarpıklık özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri

Çizelge 3.8 Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre gecikme sürelerinin Histogram alanlarının ANOVA sonuçlarını göstermektedir.

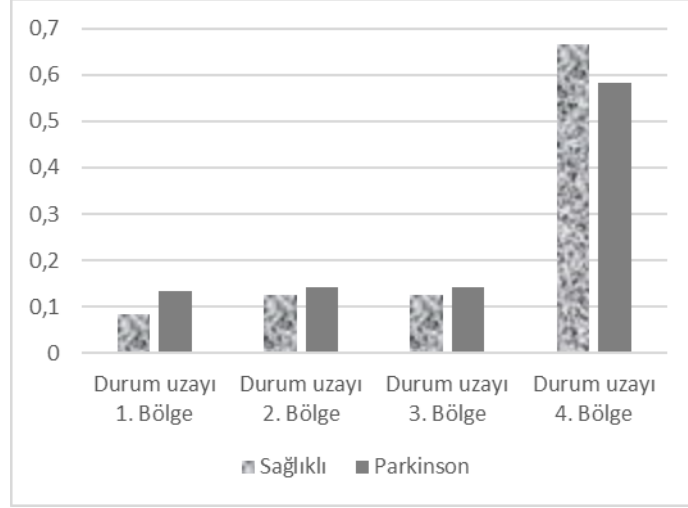
Özellik	Grup	Ortalama	Std. Sapma	Min.	Max.	T	Sig (P)
2. Alan	Normal	0,0015	0,0022	0,0000	0,0145	-3,358*	<b>0,001*</b>
	Parkinson	0,0031	0,0069	0,0000	0,0619		
18. Alan	Normal	0,0349	0,0148	0,0033	0,0756	2,476	<b>0,014</b>
	Parkinson	0,0301	0,0163	0,0000	0,0960		
34. Alan	Normal	0,0148	0,0105	0,0016	0,0434	-2,104	<b>0,036</b>
	Parkinson	0,0177	0,0118	0,0005	0,0619		
37. Alan	Normal	0,0135	0,0098	0,0013	0,0412	-2,500*	<b>0,013*</b>
	Parkinson	0,0167	0,0124	0,0000	0,0715		
39. Alan	Normal	0,0114	0,0095	0,0012	0,0472	-2,133	0,034
	Parkinson	0,0144	0,0117	0,0002	0,0597		
42. Alan	Normal	0,0093	0,0084	0,0006	0,0435	-2,887*	<b>0,004*</b>
	Parkinson	0,0125	0,0111	0,0000	0,0600		
45. Alan	Normal	0,0082	0,0079	0,0008	0,0444	-2,162*	<b>0,032*</b>
	Parkinson	0,0104	0,0094	0,0000	0,0551		



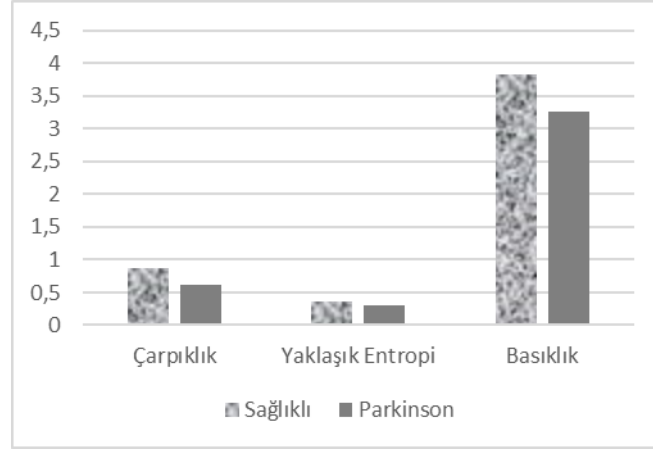
Şekil 3.17 515 kayda göre gecikme sürelerinin Histogram alanları.

Çizelge 3.9 Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre geçiş sürelerinin ANOVA sonuçlarını göstermektedir.

Özellik	Grup	Ortalama	Std. Sapma	Min.	Maks.	T	Sig (P)	
LL-RR	Normal	23,7613	27,4256	0,1738	245,863	-1,866	0,063	
	Parkinson	27,9761	22,5103	0,0051	127,475			
LR-RL	Normal	28,4095	29,0257	0,1226	154,160	-0,678	0,498	
	Parkinson	30,4960	33,1387	0,0250	274,402			
Basıklık	Normal	3,8319	1,6261	1,8401	8,7866	3,447	<b>0,001</b>	
	Parkinson	3,2715	1,7229	1,5616	14,8972			
Çarpıklık	Normal	0,8719	0,5465	-0,5427	2,1481	4,832	<b>0,000</b>	
	Parkinson	0,6116	0,5688	-0,6312	2,7722			
İlk Fark Durum Uzayı	I.	Normal	0,0824	0,0941	0,0028	0,5011	-1,457	0,146
		Parkinson	0,1322	0,1334	0,0000	0,9229		
	II.	Normal	0,1258	0,0728	0,0251	0,2731	0,403	0,687
		Parkinson	0,1426	0,0594	0,0123	0,2800		
	IV.	Normal	0,6658	0,2276	0,0916	0,9404	1,574	0,116
		Parkinson	0,5826	0,2310	0,0053	0,9700		
	III.	Normal	0,1260	0,0731	0,0251	0,2716	-1,067	0,286
		Parkinson	0,1426	0,0595	0,0123	0,2857		
Durum Uzayı	I.	Normal	0,1659	0,0832	0,0280	0,3811	-5,025*	<b>0,000*</b>
		Parkinson	0,1783	0,0904	0,0119	0,6194		
	II.	Normal	0,3108	0,0207	0,2401	0,3488	-2,503*	<b>0,013*</b>
		Parkinson	0,3100	0,0206	0,1866	0,3810		
	IV.	Normal	0,2100	0,0981	0,0563	0,4919	3,786	<b>0,000</b>
		Parkinson	0,1960	0,0927	0,0075	0,5894		
	III.	Normal	0,3131	0,0226	0,2400	0,3794	-2,474*	<b>0,014*</b>
		Parkinson	0,2240	0,0254	0,1866	0,3935		
Shannon Entropi	Normal	4,3610	0,3766	3,4511	5,0954	-0,010	0,992	
	Parkinson	4,3613	0,3367	3,2586	5,1091			
Örnek Entropi	Normal	1,7612	0,3873	0,0000	3,0802	-0,526	0,599	
	Parkinson	1,3130	2,9133	0,0000	6,5535			
Yaklaşık Entropi	Normal	0,3529	0,2125	0,0067	0,8795	2,219	<b>0,027</b>	
	Parkinson	0,3124	0,1849	-0,0104	1,2420			
Ortak Bilgi	Normal	4,2857	1,4851	2,0000	9,0000	1,232	0,218	
	Parkinson	4,1166	1,4241	2,0000	10,0000			
Yerleştirme Boyutu	Normal	4,5214	0,7342	2,0000	6,0000	0,136*	0,892*	
	Parkinson	4,5119	0,7505	2,0000	6,0000			
Ortalama	Normal	198,515	67,5331	107,3399	444,666	-0,778	0,437	
	Parkinson	204,226	69,8636	109,5790	446,574			
Medyan	Normal	178,164	69,4544	93,8000	468,800	-1,340	0,181	
	Parkinson	188,398	72,9289	93,8000	449,200			
Standart Sapma	Normal	118,709	30,7096	64,8695	190,615	0,350	0,727	
	Parkinson	117,652	27,9026	76,6677	195,054			
Varyans	Normal	15026,8	7570,83	4208,1	36334	0,533	0,594	
	Parkinson	14618,82	7127,94	5877,9	38046			



Şekil 3.18 646 kayda göre geçiş sürelerinde ayırt edici olarak bulunan durum uzayı 1., 2., 3. ve 4. Bölge nokta sayılarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



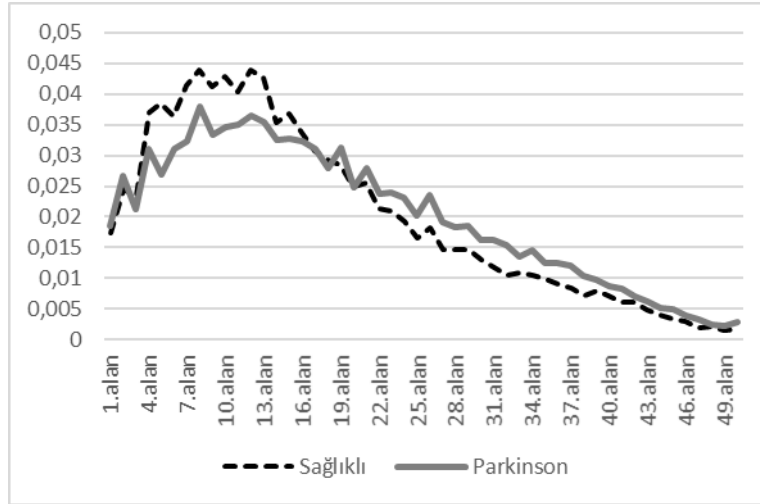
Şekil 3.19 646 kayda göre geçiş sürelerinde ayırt edici olarak bulunan Çarpıklık, Basıklık ve Yaklaşık Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.

Çizelge 3.10 Sağlıklı ve Parkinson hastası katılımcıların 646 kayda göre geçiş sürelerinin Histogram alanlarının ANOVA sonuçlarını göstermektedir.

Özellik	Grup	Ortalama	Std. Sapma	Minimum	Maksimum	T	Sig (P)
5. Alan	Normal	0,0384	0,0316	0,0000	0,2273	1,930	<b>0,000</b>
	Parkinson	0,0268	0,213	0,0000	0,1094		
6. Alan	Normal	0,0364	0,0236	0,0000	0,1069	2,407	<b>0,017</b>
	Parkinson	0,0311	0,0216	0,0000	0,1139		
7. Alan	Normal	0,0413	0,0222	0,0000	0,1046	4,326	<b>0,000</b>
	Parkinson	0,0323	0,0219	0,0000	0,0999		
8. Alan	Normal	0,0439	0,0230	0,0000	0,1188	2,639	<b>0,009</b>
	Parkinson	0,0380	0,0234	0,0000	0,1691		
9. Alan	Normal	0,0412	0,0211	0,0000	0,0895	3,945	<b>0,000</b>
	Parkinson	0,0334	0,0207	0,0000	0,1026		
10. Alan	Normal	0,0404	0,0174	0,0000	0,0935	4,362	<b>0,000</b>
	Parkinson	0,0350	0,0203	0,0000	0,1429		
11. Alan	Normal	0,0404	0,0189	0,0000	0,0906	3,089	<b>0,002</b>
	Parkinson	0,0350	0,0182	0,0000	0,1397		
12. Alan	Normal	0,0439	0,0176	0,0000	0,0921	4,464	<b>0,000</b>
	Parkinson	0,0364	0,0175	0,0000	0,0905		
13. Alan	Normal	0,0425	0,0166	0,0000	0,0872	4,450	<b>0,000</b>
	Parkinson	0,0354	0,0168	0,0000	0,1012		
15. Alan	Normal	0,0367	0,0158	0,0000	0,1250	2,697	<b>0,007</b>
	Parkinson	0,0327	0,0157	0,0000	0,0972		
22. Alan	Normal	0,0213	0,0129	0,0024	0,0723	-2,113	<b>0,035</b>
	Parkinson	0,0238	0,0128	0,0000	0,0769		
23. Alan	Normal	0,0209	0,0130	0,0000	0,0621	-2,516	<b>0,012</b>
	Parkinson	0,0240	0,0133	0,0000	0,0800		
24. Alan	Normal	0,0193	0,0117	0,0000	0,0597	-3,531	<b>0,000</b>
	Parkinson	0,0232	0,0117	0,0000	0,0800		
25. Alan	Normal	0,0166	0,0102	0,0000	0,0507	-3,391	<b>0,001</b>
	Parkinson	0,0201	0,0129	0,0000	0,0784		
26. Alan	Normal	0,0181	0,0122	0,0021	0,0759	-4,250	<b>0,000</b>
	Parkinson	0,0235	0,0134	0,0000	0,1034		
27. Alan	Normal	0,0147	0,0106	0,0000	0,0455	-3,977	<b>0,000</b>
	Parkinson	0,0192	0,0122	0,0000	0,0769		
28. Alan	Normal	0,0146	0,0107	0,0000	0,0581	-3,118	<b>0,002</b>
	Parkinson	0,0184	0,0132	0,0000	0,1034		
29. Alan	Normal	0,0147	0,0110	0,0000	0,0503	-3,337	<b>0,001</b>
	Parkinson	0,0185	0,0124	0,0000	0,0857		
30. Alan	Normal	0,0129	0,0097	0,0000	0,0410	-3,380	<b>0,001</b>
	Parkinson	0,0163	0,0123	0,0000	0,0617		
31. Alan	Normal	0,0117	0,0093	0,0000	0,0406	-4,691	<b>0,000</b>
	Parkinson	0,0161	0,0121	0,0000	0,0761		
32. Alan	Normal	0,0105	0,0090	0,0000	0,0405	-5,334	<b>0,000</b>
	Parkinson	0,0154	0,0114	0,0000	0,0554		
33. Alan	Normal	0,0109	0,0093	0,0000	0,0414	-2,809	<b>0,005</b>
	Parkinson	0,0136	0,0118	0,0000	0,0598		
34. Alan	Normal	0,0104	0,0092	0,0000	0,0404	-4,405	<b>0,000</b>
	Parkinson	0,0144	0,0111	0,0000	0,0555		

Çizelge 3.10 devam ediyor.

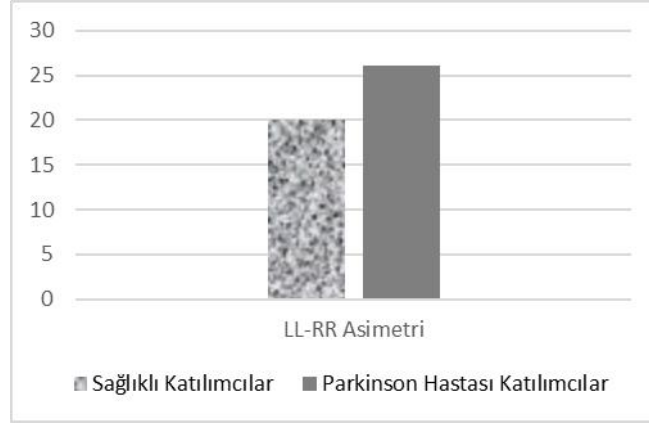
35. Alan	Normal	0,0099	0,0095	0,0000	0,0446	-2,621	<b>0,009</b>
	Parkinson	0,0125	0,0118	0,0000	0,1207		
36. Alan	Normal	0,0091	0,0087	0,0000	0,0426	-3,714	<b>0,000</b>
	Parkinson	0,0125	0,0116	0,0000	0,0725		
37. Alan	Normal	0,0085	0,0082	0,0000	0,0384	-4,085	<b>0,000</b>
	Parkinson	0,0120	0,0120	0,0000	0,1062		
38. Alan	Normal	0,0072	0,0071	0,0000	0,0342	-4,066	<b>0,000</b>
	Parkinson	0,0103	0,0108	0,0000	0,1129		
41. Alan	Normal	0,0062	0,0075	0,0000	0,0483	-2,809	<b>0,005</b>
	Parkinson	0,0083	0,0094	0,0000	0,0667		
45. Alan	Normal	0,0000	0,0060	0,0000	0,0000	-2,375	<b>0,018</b>
	Parkinson	0,0100	0,0080	0,0000	0,0000		
47. Alan	Normal	0,0020	0,0037	0,0000	0,0200	-3,141	<b>0,002</b>
	Parkinson	0,0033	0,0053	0,0000	0,0400		
50. Alan	Normal	0,0000	0,0000	0,0000	0,0000	-3,555	<b>0,000</b>
	Parkinson	0,0000	0,0000	0,0000	0,0000		



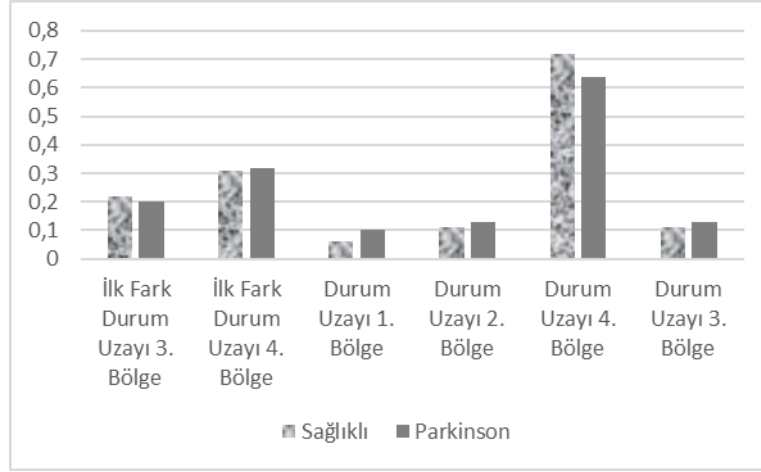
Şekil 3.20 646 kayda göre geçiş sürelerinin Histogram alanları.

Çizelge 3.11 Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre geçiş sürelerinin ANOVA sonuçları.

Özellik	Grup	Ortalama	Std. Sapma	Min.	Maks.	T	Sig (P)	
LL-RR	Normal	20,0496	18,1068	0,1738	101,5815	-3,046*	<b>0,003*</b>	
	Parkinson	26,0569	20,3811	0,0051	100,3668			
LR-RL	Normal	23,6169	25,2025	0,1226	154,1600	-0,701	0,483	
	Parkinson	25,4246	24,1091	0,2217	155,0897			
Basıklık	Normal	4,0964	1,6586	1,9516	8,7866	3,994*	<b>0,000*</b>	
	Parkinson	3,3880	1,7364	1,6721	14,8972			
Çarpıklık	Normal	0,9514	0,5356	-0,5427	2,1481	4,668	<b>0,000</b>	
	Parkinson	0,6868	0,5358	-0,6312	2,7722			
İlk Fark Durum Uzaı	I.	Normal	0,1637	0,0850	0,0280	0,3613	-1,225	0,221
		Parkinson	0,1746	0,0841	0,0171	0,4283		
	II.	Normal	0,3076	0,0190	0,2401	0,3429	-1,034	0,302
		Parkinson	0,3096	0,0180	0,1965	0,3470		
	IV.	Normal	0,2186	0,0992	0,0596	0,4919	2,002	<b>0,046</b>
		Parkinson	0,1991	0,0900	0,0451	0,5894		
III.	Normal	0,3101	0,0218	0,2400	0,3794	-2,559	<b>0,011</b>	
	Parkinson	0,3167	0,0247	0,1971	0,3935			
Durum Uzaı	I.	Normal	0,0633	0,0756	0,0028	0,5011	-4,497*	<b>0,000*</b>
		Parkinson	0,1026	0,1032	0,0009	0,5445		
	II.	Normal	0,1096	0,0990	0,0251	0,2599	-3,131*	<b>0,002*</b>
		Parkinson	0,1309	0,0664	0,0123	0,2566		
	IV.	Normal	0,7176	0,0549	0,0916	0,9404	3,878	<b>0,000</b>
		Parkinson	0,6359	0,1992	0,1081	0,9700		
III.	Normal	0,1096	0,0665	0,0251	0,2605	-3,125*	<b>0,002*</b>	
	Parkinson	0,1308	0,0549	0,0123	0,2566			
Shannon Entropi	Normal	4,3348	0,3799	3,4511	5,0954	0,925	0,356	
	Parkinson	4,3601	0,3188	3,6495	5,1091			
Örnek Entropi	Normal	1,7482	0,2852	0,1670	2,3905	-2,276	<b>0,023</b>	
	Parkinson	1,8159	0,2802	0,4451	2,8207			
Yaklaşık Entropi	Normal	0,3964	0,2035	0,0566	0,8795	2,655	<b>0,008</b>	
	Parkinson	0,3448	0,1773	0,0113	1,2420			
Ortak Bilgi	Normal	4,5391	1,4406	2,0000	9,0000	1,358	0,175	
	Parkinson	4,3425	1,3471	2,0000	9,0000			
Yerleşirme Boyutu	Normal	4,7391	0,5475	3,0000	6,0000	-0,933	0,351	
	Parkinson	4,7875	0,4721	3,0000	6,0000			
Ortalama	Normal	198,5159	67,5331	107,33	444,6662	-0,778	0,437	
	Parkinson	204,2260	69,8635	109,57	446,5748			
Medyan	Normal	178,1648	69,4544	93,800	468,8000	-1,340	0,181	
	Parkinson	188,3989	72,9289	93,800	449,2000			
Standart Sapma	Normal	118,7094	30,7096	64,869	190,6159	0,350	0,727	
	Parkinson	117,6529	27,9026	76,667	195,0543			
Varyans	Normal	15026,8	7570,8	4208,1	36334	0,533	0,594	
	Parkinson	14618,8	7127,9	5877,9	38046			

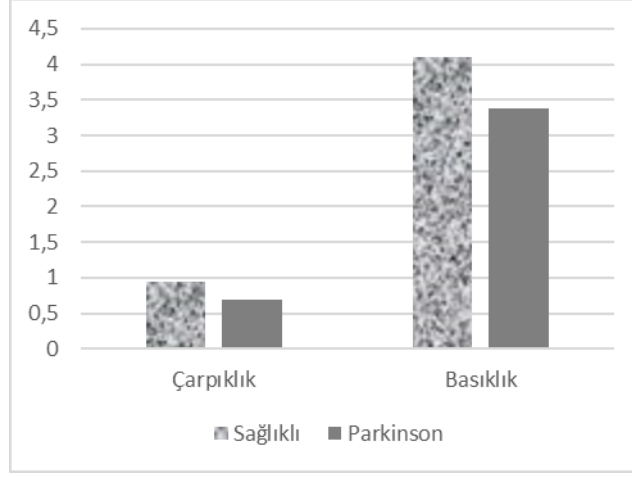


Şekil 3.21 515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan LL-RR Asimetri özelliğinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.

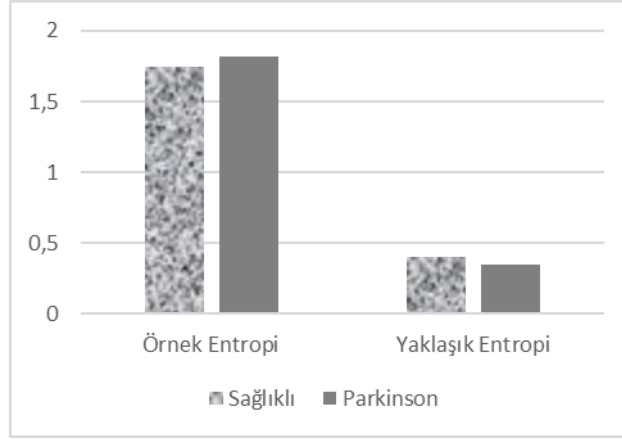


Şekil 3.22 515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan ilk fark durum uzayı 1., 2., 3. ve 4. bölge nokta sayılarının sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.





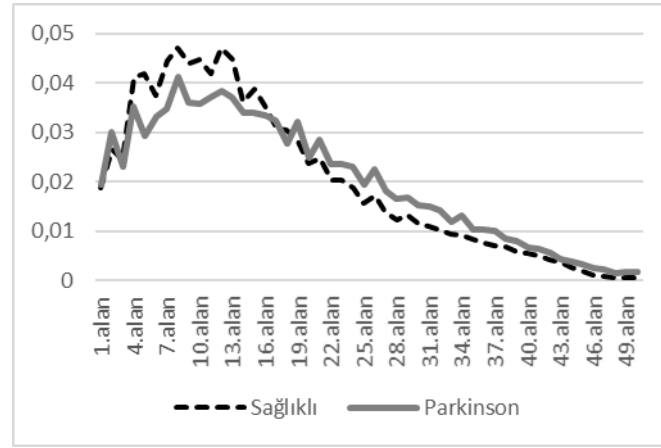
Şekil 3.23 515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan Çarpıklık ve Basıklık özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.



Şekil 3.24 515 kayda göre geçiş sürelerinde ayırt edici olarak bulunan Örnek Entropi ve Yaklaşık Entropi özelliklerinin sağlıklı ve Parkinson hastası bireyler için ortalama değerleri.

Çizelge 3.12 Sağlıklı ve Parkinson hastası katılımcıların 515 kayda göre geçiş sürelerinin Histogram alanlarının ANOVA sonuçları.

Özellik	Grup	Ortalama	Std. Sapma	Min.	Max.	t	Sig (P)
5. Alan	Normal	0,0363	0,0274	0,0010	0,1387	2,169	<b>0,031</b>
	Parkinson	0,0300	0,0221	0,0000	0,1094		
7. Alan	Normal	0,0411	0,0213	0,0027	0,1047	2,147	<b>0,032</b>
	Parkinson	0,0353	0,0221	0,0000	0,0976		
10. Alan	Normal	0,0429	0,0178	0,0018	0,0935	2,604	<b>0,010</b>
	Parkinson	0,0368	0,0193	0,0000	0,0908		
12. Alan	Normal	0,0452	0,0172	0,0030	0,0921	3,657	<b>0,000</b>
	Parkinson	0,0377	0,0166	0,0000	0,0905		
13. Alan	Normal	0,0437	0,0166	0,0024	0,0873	2,913	<b>0,004</b>
	Parkinson	0,0381	0,0153	0,0000	0,0835		
15. Alan	Normal	0,0390	0,0137	0,0102	0,0879	2,622	<b>0,009</b>
	Parkinson	0,0344	0,0143	0,0000	0,0853		
16. Alan	Normal	0,0356	0,0144	0,0053	0,0988	2,051	<b>0,041</b>
	Parkinson	0,0324	0,0119	0,0034	0,0790		
18. Alan	Normal	0,0320	0,0141	0,0058	0,0922	2,783	<b>0,006</b>
	Parkinson	0,0276	0,0127	0,0000	0,0762		
28. Alan	Normal	0,0140	0,0009	0,0022	0,0455	-2,161	<b>0,031</b>
	Parkinson	0,0170	0,0006	0,0000	0,0763		
31. Alan	Normal	0,0123	0,0096	0,0012	0,0406	-2,057	<b>0,040</b>
	Parkinson	0,0150	0,0112	0,0012	0,0658		
32. Alan	Normal	0,0120	0,0094	0,0017	0,0405	-2,087	<b>0,038</b>
	Parkinson	0,0146	0,0103	0,0000	0,0540		
34. Alan	Normal	0,0105	0,0085	0,0011	0,0404	-2,637	<b>0,009</b>
	Parkinson	0,0136	0,0099	0,0000	0,0555		
36. Alan	Normal	0,0089	0,0076	0,0008	0,0426	-1,976	<b>0,049</b>
	Parkinson	0,01120	0,0099	0,0000	0,0591		
37. Alan	Normal	0,0083	0,0077	0,0010	0,0384	-2,142	<b>0,033</b>
	Parkinson	0,0108	0,0100	0,0004	0,0722		
44. Alan	Normal	0,0032	0,0043	0,0000	0,0280	-2,099*	<b>0,037*</b>
	Parkinson	0,0044	0,0062	0,0000	0,0610		
45. Alan	Normal	0,0022	0,0035	0,0000	0,0280	-2,967*	<b>0,003*</b>
	Parkinson	0,0037	0,0056	0,0000	0,0297		
46. Alan	Normal	0,0014	0,0025	0,0000	0,0208	-3,912*	<b>0,000*</b>
	Parkinson	0,0027	0,0037	0,0000	0,0211		
47. Alan	Normal	0,0011	0,0018	0,0000	0,0100	-4,619*	<b>0,000*</b>
	Parkinson	0,0025	0,0037	0,0000	0,0300		
48. Alan	Normal	0,0006	0,0008	0,0000	0,0039	-5,249*	<b>0,000*</b>
	Parkinson	0,0015	0,0025	0,0000	0,0189		
49. Alan	Normal	0,0006	0,0008	0,0000	0,0041	-6,070*	<b>0,000*</b>
	Parkinson	0,0013	0,0022	0,0000	0,0104		
50. Alan	Normal	0,0007	0,00066	0,0000	0,0033	-6,240*	<b>0,000*</b>
	Parkinson	0,0016	0,0020	0,0001	0,0116		



Şekil 3.25 515 kayda göre geçiş sürelerinin Histogram alanları.

### 3.2 Sınıflayıcı Sonuçları

İstatistiksel analiz sonrası elde ettiğimiz sonuçlardan yola çıkarak farklı değerlendirmeler yapıldı. Yapılan değerlendirmeler, test ve eğitim oranının farklı iki durumunda elde edilmiştir. Bir tanesinde veri setinin 50%'si eğitim, geriye kalan 50%'si ise test için ele alınırken, diğer durumda ise veri setinin 70%'i eğitim, geriye kalan 30%'luk kısmı ise test için ele alınmıştır. İlk adımda istatistiksel analiz sonucunda P değeri 0,05'ten küçük olan tüm özellikler tekerli halde iki farklı sınıflayıcının girişine uygulanmış ve elde edilen tüm sonuçlar listelenmiştir. İkinci durumda P değeri en düşük çıkan beş özellik belirlenmiş ve bu beş özellik iki farklı sınıflayıcının girişine aynı anda uygulanmış ve elde edilen sonuçlar listelenmiştir. Üçüncü durumda P değeri 0,001'e eşit ve küçük çıkan tüm özellikler aynı anda iki farklı sınıflayıcının girişine uygulanmış ve elde edilen sonuçlar listelenmiştir. Dördüncü ve son durumda ise P değeri 0,000 çıkan tüm özelliklere RO algoritması uygulanmış, içerisinde hangi özelliğin sırası ile ne kadar değerli olduğu belirlenmiş ve en değersiz olarak belirlenen özellikler sırası ile tüm özellikler içerisinde teker teker çıkartılarak sınıflayıcı sonuçları listelenmiştir. 646 ve 515 kayıt için elde edilen sonuçlar ayrı ayrı verilmiştir.

### 3.2.1 Destek Vektör Makinesi Sınıflayıcısı Sonuçları

Aşağıdaki çizelgelerde hem 646 kayda hem de 515 kayda göre en ayırt edici olarak bulunan 5 özellik tekli ve beşerli şekilde Destek Vektör Makinesi sınıflayıcısı girişine verilerek sonuçları listelenmiştir.

Çizelge 3.13 646 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde Destek Vektör Makinesi sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi	Yaklaşık Entropi	78,421	77,925	78,650	77,371
	İlk Fark Durum Uzayı IV. Bölge	79,040	77,523	78,384	78,165
	LL-RR Asimetri	78,390	74,582	78,385	75,360
	LR-RL Asimetri	76,843	75,944	78,320	75,670
	Ortalama	77,492	69,845	77,256	69,742
	Medyan	78,049	78,421	78,340	78,092
	Shannon Entropi	79,040	77,430	79,982	76,443
	Standart Sapma	76,904	68,049	77,787	70,309
	Varyans	78,452	78,111	78,849	76,958
Gecikme Süresi	Basıklık	77,770	77,647	78,296	77,010
	Çarpıklık	78,544	77,709	77,721	79,329
	Durum Uzayı II. Bölge	77,956	78,699	78,075	78,917
	Durum Uzayı I. Bölge	78,761	77,894	78,053	77,216
	Durum Uzayı IV. Bölge	78,792	77,863	78,163	78,711
	Durum Uzayı III. Bölge	78,637	78,018	77,986	<b>79,123</b>
	Ortak Bilgi	77,554	79,040	78,451	78,041
	Yaklaşık Entropi	78,204	78,452	78,827	77,164
Geçiş Süresi	Yaklaşık Entropi	78,699	77,925	78,119	78,814
	Basıklık	77,678	77,337	77,831	77,628
	Çarpıklık	78,513	77,956	78,630	78,556
	Durum Uzayı II. Bölge	77,131	78,210	77,783	77,402
	Durum Uzayı I. Bölge	77,015	78,326	77,922	77,077
	Durum Uzayı IV. Bölge	77,054	78,015	78,088	76,233
	Durum Uzayı III. Bölge	78,410	76,848	77,285	78,571

Çizelge 3.14 515 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde Destek Vektör Makinesi sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi	Shannon Entropi	77,713	77,159	78,670	74,480
	Yaklaşık Entropi	78,333	77,003	77,700	77,597
	İlk Fark Durum Uzayı IV.Bölge	77,713	77,626	78,476	75,779
	LL-RR Asimetri	78,255	73,151	77,867	74,545
	LR-RL Asimetri	77,131	73,229	77,950	74,025
	Ortalama	76,860	67,159	76,675	68,896
	Medyan	78,294	77,003	78,116	76,623
	Standart Sapma	78,992	66,614	76,925	71,753
Gecikme Süresi	Varyans	78,604	76,575	77,063	<b>78,506</b>
	Basıklık	77,558	75,797	77,340	76,883
	Çarpıklık	78,139	77,003	77,396	77,857
	Durum Uzayı II. Bölge	78,410	76,926	77,977	76,948
	Durum Uzayı IV Bölge.	77,248	77,704	77,451	78,181
Geçiş Süresi	Durum Uzayı III. Bölge	77,209	77,859	78,116	76,623
	Basıklık	76,976	76,264	77,146	75,844
	Çarpıklık	77,286	77,898	77,922	77,012
	Durum Uzayı II. Bölge	77,374	77,665	77,617	77,792
	Durum Uzayı I. Bölge	77,015	78,326	77,950	77,012
	Durum Uzayı IV. Bölge	76,892	78,171	77,839	77,272

Çizelge 3.15 646 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde Destek Vektör Makinesi sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi, Basılı Tutma Süresi, Basılı Tutma Süresi, Gecikme Süresi, Geçiş Süresi	LL-RR Asimetri, LR-RL Asimetri, Varyans, Çarpıklık, Çarpıklık	78,390	78,266	78,325	<b>78,313</b>

Çizelge 3.16 515 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde Destek Vektör Makinesi sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi, Basılı Tutma Süresi, Basılı Tutma Süresi, Gecikme Süresi, Geçiş Süresi	Yaklaşık Entropi, LL-RR Asimetri, LR-RL Asimetri, Çarpıklık, Basıklık	77,596	<b>77,237</b>	77,451	76,883

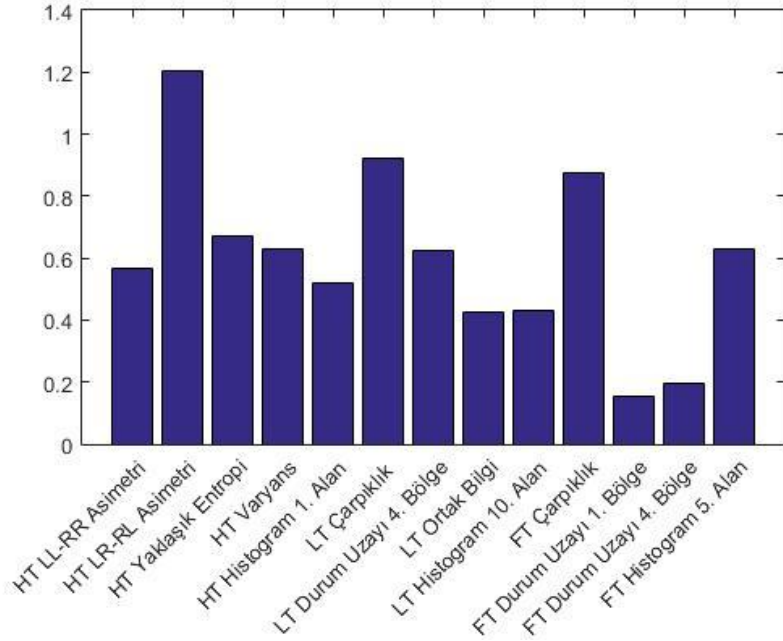
Çizelge 3.17 646 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte Destek Vektör Makinesi sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Özellikler	50-50		30-70	
	Train	Test	Train	Test
$P \leq 0.05$ olan tüm özellikler	78,482	<b>78,713</b>	78,384	78,195

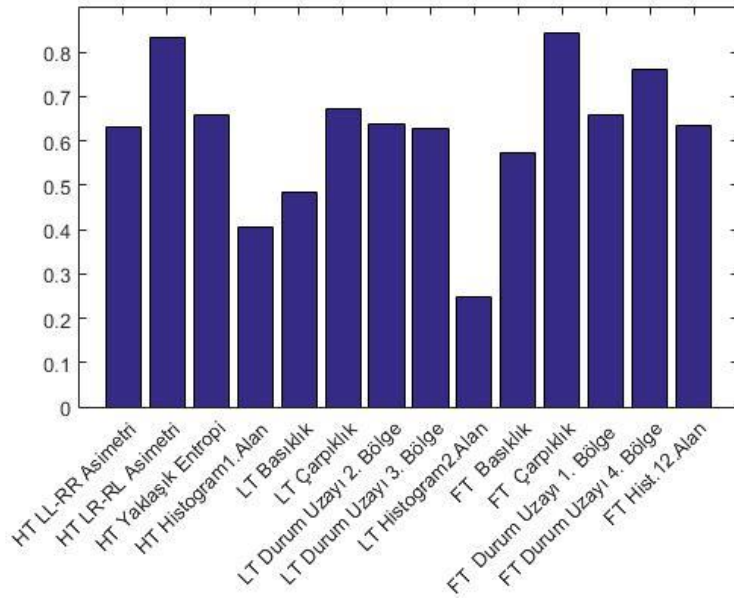
Çizelge 3.18 515 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte Destek Vektör Makinesi sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Özellikler	50-50		30-70	
	Train	Test	Train	Test
$P \leq 0.05$ olan tüm özellikler	78,482	76,860	77,617	<b>77,792</b>

Rasgele Orman algoritması ile özellik seçimi yapılmış ve hangi özelliğin ne kadar değerli olduğu belirlenmiştir. Bu algoritma  $P \leq 0,001$  çıkan özelliklere uygulanmıştır. Şekil 3.26 ve Şekil 3.27’de bu özelliklerin hangileri olduğu ve değer dereceleri gösterilmiştir.



Şekil 3.26 646 kayda göre  $P \leq 0,001$  çıkan 13 özelliğin önem derecelerini göstermektedir.



Şekil 3.27 515 kayda göre  $P \leq 0,001$  çıkan 14 özelliğin önem derecelerini göstermektedir.

Şekil 3.26 ve 3.27'de değerlendirilen özellikler birlikte sınıflayıcıya sokulmuş ve en değersiz özellikler sırası ile tümünün içerisinde çıkarılarak sınıflayıcı doğruluk oranları tekrar test edilmiştir. Elde edilen bu sonuçlar Çizelge 3.19 ve 3.20'de gösterilmiştir.

Çizelge 3.19 Rasgele Orman algoritması ile 646 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen Destek Vektör Makinesi sınıflayıcısı doğruluk oranları.

Sırası ile tümünden çıkarılan özellikler	50-50		30-70	
	Train	Test	Train	Test
Geçiş Süresi Durum Uzayı IV. Bölge	78,235	78,421	78,075	78,917
Geçiş Süresi Durum Uzayı I. Bölge	79,133	77,523	78,053	78,969
Gecikme Süresi Histogram 10. Alan	78,292	78,359	77,986	<b>79,123</b>
Basılı Tutma Histogram 1. Alan	79,504	77,151	78,411	78,762
Gecikme Süresi Ortak Bilgi	79,164	77,492	78,119	78,814
Basılı Tutma Süresi Yaklaşık Entropi	78,111	78,544	78,584	77,731
Basılı Tutma Süresi LL-RR Asimetri	78,854	77,801	77,942	79,226
Basılı Tutma Süresi Varyans	77,956	67,523	78,561	67,938
Gecikme Süresi Çarpıklık	77,987	66,965	77,610	66,649
Gecikme Süresi Durum Uzayı IV. Bölge	76,637	64,860	77,898	66,082
Geçiş Süresi Histogram 5. Alan	77,306	65,387	77,544	66,237
Geçiş Süresi Çarpıklık	78,761	74,582	77,654	77,835



Çizelge 3.20 Rasgele Orman algoritması ile 515 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen Destek Vektör Makinesi sınıflayıcısı doğruluk oranları.

Sırası ile tümünden çıkarılan özellikler	50-50		30-70	
	Train	Test	Train	Test
Gecikme Süresi Histogram 2. Alan	77,170	<b>78,054</b>	77,423	77,792
Basılı Tutma Histogram 1. Alan	78,178	76,926	77,811	77,142
Geçiş Süresi Basıklık	77,286	77,315	77,950	75,584
Gecikme Süresi Basıklık	77,984	75,603	77,423	75,584
Geçiş Süresi Histogram 12. Alan	79,108	74,747	77,534	75,909
Gecikme Süresi Durum Uzayı II. Bölge	78,217	75,914	77,506	76,168
Gecikme Süresi Durum Uzayı III. Bölge	77,364	76,186	78,005	75,122
Basılı Tutma Süresi LL-RR Asimetri	76,899	71,089	77,562	71,103
Geçiş Süresi Durum Uzayı I. Bölge	77,364	70,466	77,479	70,324
Gecikme Süresi Çarpıklık	77,635	67,821	77,063	68,441
Geçiş Süresi Durum Uzayı IV. Bölge	77,829	65,058	78,171	65,064
Basılı Tutma Süresi Yaklaşık Entropi	77,286	63,813	76,565	66,363
Geçiş Süresi Çarpıklık	77,674	73,346	77,562	74,935

### 3.2.2 k En Yakın Komşu Sınıflayıcısı Sonuçları

Aşağıdaki çizelgelerde hem 646 kayda hem de 515 kayda göre en ayırt edici olarak bulunan 5 özellik tekli ve beşerli şekilde k En Yakın Komşu sınıflayıcısına sokularak sonuçları listelenmiştir.

Çizelge 3.21 646 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde k En Yakın Komşu sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi	Yaklaşık Entropi	78,988	75,098	79,550	74,999
	İlk Fark Durum Uzayı IV. Bölge	78,811	76,148	79,119	75,927
	LL-RR Asimetri	79,865	75,562	80,346	76,048
	LR-RL Asimetri	79,937	75,550	79,528	75,893
	Ortalama	80,718	75,801	80,663	76,072
	Medyan	80,798	77,108	80,884	76,885
	Shannon Entropi	81,166	77,347	81,209	77,440
	Standart Sapma	79,484	76,388	79,653	75,653
	Varyans	79,640	76,031	79,318	76,484
Gecikme Süresi	Basıklık	80,103	77,069	79,948	77,508
	Çarpıklık	82,002	78,214	82,079	78,058
	Durum Uzayı I. Bölge	80,346	75,739	80,083	76,535
	Durum Uzayı II. Bölge	79,715	75,690	79,389	75,278
	Durum Uzayı IV. Bölge	81,919	76,953	79,884	77,381
	Durum Uzayı III. Bölge	79,680	75,869	79,668	75,164
	Yaklaşık Entropi	78,874	75,562	79,011	75,124
	Ortak Bilgi	77,389	75,737	78,098	76,529
Geçiş Süresi	Yaklaşık Entropi	79,133	75,593	79,742	75,051
	Basıklık	79,277	76,088	80,117	75,893
	Çarpıklık	79,607	75,892	80,147	76,065
	Durum Uzayı I. Bölge	80,689	76,732	80,759	75,584
	Durum Uzayı II. Bölge	81,542	76,910	81,489	76,766
	Durum Uzayı IV. Bölge	81,765	77,525	81,983	<b>78,155</b>
	Durum Uzayı III. Bölge	81,325	76,179	81,639	76,935

Çizelge 3.22 515 kayda göre ayırt edici olarak bulunan tüm özelliklerin tekerli halde k En Yakın Komşu sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi	Yaklaşık Entropi	79,212	76,122	79,298	74,696
	İlk Fark Durum Uzayı IV. Bölge	78,346	74,811	78,061	75,865
	LL-RR Asimetri	79,638	74,850	79,870	74,588
	LR-RL Asimetri	79,44	75,214	79,584	74,155
	Ortalama	79,160	74,448	79,778	75,130
	Medyan	79,147	76,718	78,688	76,926
	Shannon Entropi	81,408	76,978	81,265	77,489
	Standart Sapma	82,080	79,623	83,148	<b>79,999</b>
	Varyans	83,385	78,923	82,825	79,598
Gecikme Süresi	Basıklık	79,987	76,368	79,713	77,554
	Çarpıklık	81,705	77,107	81,689	77,424
	Durum Uzayı II. Bölge	81,240	76,796	81,541	76,515
	Durum Uzayı IV. Bölge	81,473	77,717	82,031	78,354
	Durum Uzayı III. Bölge	81,227	76,848	81,865	76,233
Geçiş Süresi	Basıklık	80,969	75,759	80,554	75,887
	Çarpıklık	79,729	75,006	79,520	75,952
	Durum Uzayı I. Bölge	80,813	76,069	80,674	75,671
	Durum Uzayı II. Bölge	81,266	76,459	81,643	76,471
	Durum Uzayı IV. Bölge	82,092	76,653	81,985	77,900

Çizelge 3.23 646 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde k En Yakın Komşu sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi, Basılı Tutma Süresi, Basılı Tutma Süresi, Gecikme Süresi, Geçiş Süresi	LL-RR Asimetri, LR-RL Asimetri, Varyans, Çarpıklık, Çarpıklık	79,845	<b>76,140</b>	79,727	75,601

Çizelge 3.24 515 kayda göre en ayırt edici olarak bulunan 5 özelliğin beşerli şekilde k En Yakın Komşu sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Zaman Serisi	Özellik	50-50		30-70	
		Train	Test	Train	Test
Basılı Tutma Süresi, Basılı Tutma Süresi, Basılı Tutma Süresi, Gecikme Süresi, Geçiş Süresi	Yaklaşık Entropi, LL-RR Asimetri, LR-RL Asimetri, Çarpıklık, Basıklık	82,920	79,221	83,241	<b>80,151</b>

Çizelge 3.25 646 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte k En Yakın Komşu sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

Özellikler	50-50		30-70	
	Train	Test	Train	Test
$P \leq 0.05$ olan tüm özellikler	81,424	79,131	81,969	<b>79,587</b>

Çizelge 3.26 515 kayda göre ayırt edici olarak belirlenen tüm özellikler hepsi birlikte k En Yakın Komşu sınıflayıcısı girişlerine verilerek elde edilen sonuçları.

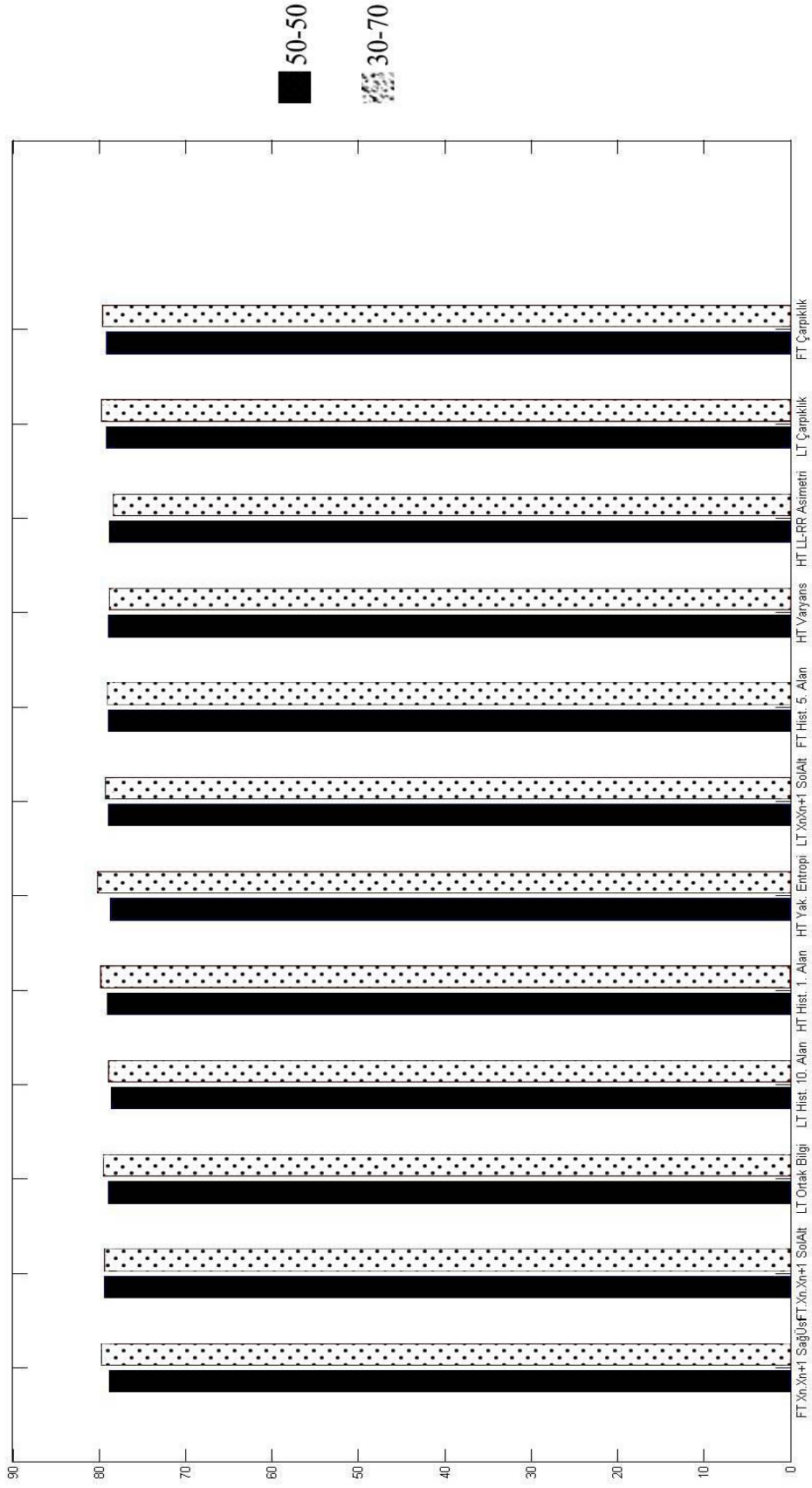
Özellikler	50-50		30-70	
	Train	Test	Train	Test
$P \leq 0.05$ olan tüm özellikler	84,147	78,677	83,074	<b>80,129</b>

Çizelge 3.27 Rasgele Orman algoritması ile 646 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.

Sırası ile tümünden çıkarılan özellikler	50-50		30-70	
	Train	Test	Train	Test
Geçiş Süresi Durum Uzayı IV. Bölge	83,4365	78,8235	83,3850	79,6907
Geçiş Süresi Durum Uzayı I. Bölge	84,3034	79,4118	84,0929	79,3814
Gecikme Süresi Histogram 10. Alan	83,3437	78,9783	83,6504	79,4845
Basılı Tutma Histogram 1. Alan	83,4985	78,6378	83,6947	78,9175
Gecikme Süresi Ortak Bilgi	82,8793	79,0402	83,9381	79,8454
Basılı Tutma Süresi Yaklaşık Entropi	83,1889	78,7616	83,7832	<b>80,1546</b>

Çizelge 3.27 devam ediyor.

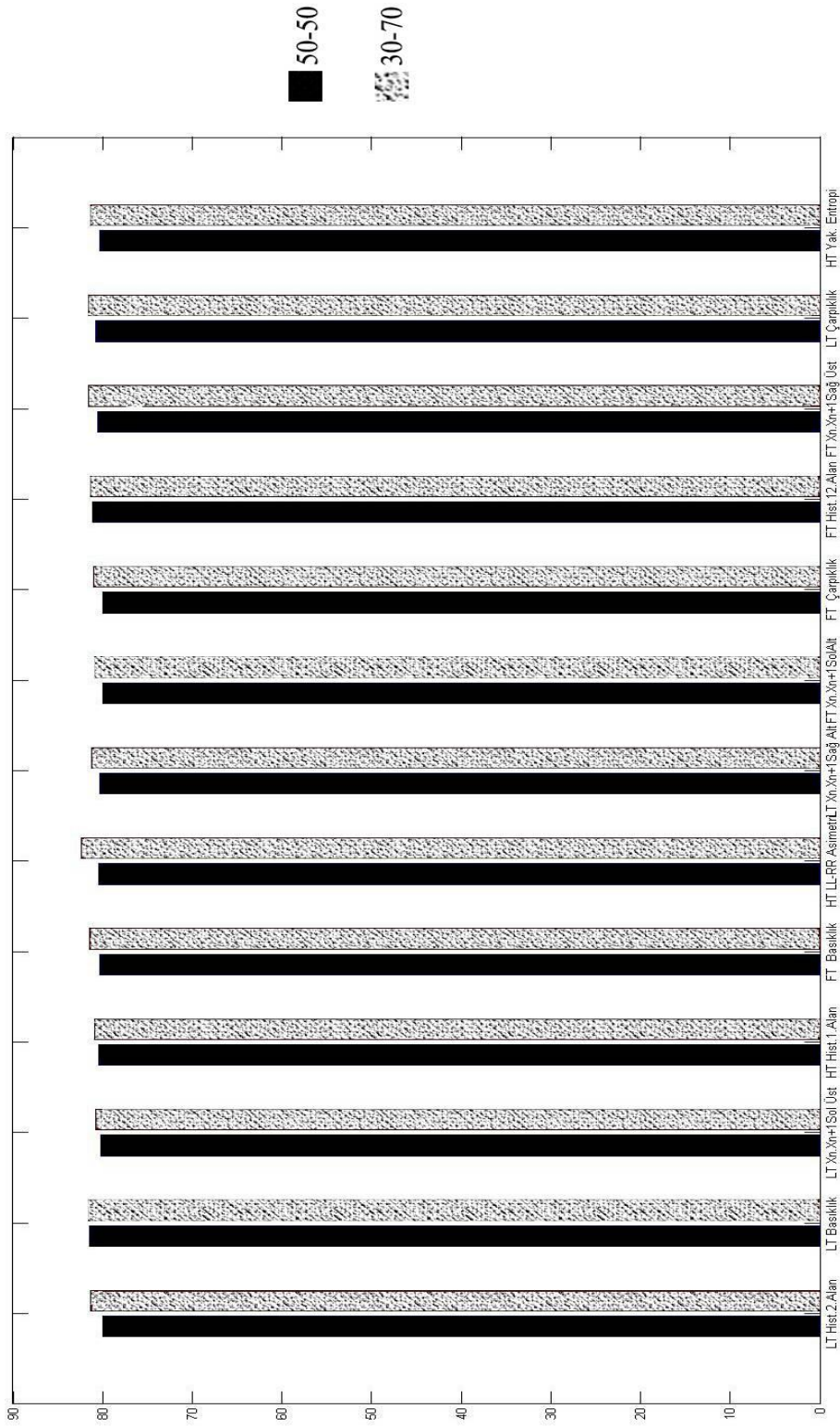
Basılı Tutma Süresi LL-RR Asimetri	83,9938	78,8854	83,4513	79,2784
Basılı Tutma Süresi Varyans	84,3344	78,9474	83,8496	79,0722
Gecikme Süresi Çarpıklık	84,0557	78,8854	83,8938	78,8144
Gecikme Süresi Durum Uzayı IV. Bölge	83,3437	78,7926	84,2257	78,4021
Geçiş Süresi Histogram 5. Alan	83,5913	79,1331	83,8274	79,7423
Geçiş Süresi Çarpıklık	84,0248	79,131	83,2301	79,6392



Şekil 3.28 Rasgele Orman algoritması ile 646 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığı elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.

Çizelge 3.28 Rasgele Orman algoritması ile 515 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.

Sırası ile tümünden çıkarılan özellikler	50-50		30-70	
	Train	Test	Train	Test
Gecikme Süresi Histogram 2. Alan	87,5581	80	87,2299	81,2987
Basılı Tutma Histogram 1. Alan	87,0155	81,4786	87,4238	81,5584
Geçiş Süresi Basıklık	87,2480	80,2334	87,4238	80,7792
Gecikme Süresi Basıklık	86,4341	80,4280	87,1745	80,8441
Geçiş Süresi Histogram 12. Alan	87,0930	80,3112	87,3684	81,4285
Gecikme Süresi Durum Uzayı II. Bölge	87,2093	80,3891	86,6482	<b>82,4026</b>
Gecikme Süresi Durum Uzayı III. Bölge	86,5503	80,3112	86,9806	81,2337
Basılı Tutma Süresi LL-RR Asimetri	87,3643	80	86,7867	80,8441
Geçiş Süresi Durum Uzayı I. Bölge	87,2093	79,9610	87,7008	81,0389
Gecikme Süresi Çarpıklık	87,7441	81,1673	87,2292	81,2987
Geçiş Süresi Durum Uzayı IV. Bölge	87,3255	80,5863	87,0914	81,6233
Basılı Tutma Süresi Yaklaşık Entropi	87,6744	80,7781	86,5651	81,6233
Geçiş Süresi Çarpıklık	85,9689	80,3509	87,3407	81,3636



Şekil 3.29 Rasgele Orman algoritması ile 515 kayda göre değer sıralaması belirlenen özelliklerin sırayla tümünden çıkarıldığında elde edilen k En Yakın Komşu sınıflayıcısı doğruluk oranları.

646 kayıt ve 515 kayıt için en yüksek doğruluk oranlarına sahip iki sonucun konfüzyon matrisleri Çizelge 3.29 ve 3.30'da gösterilmiştir.



Çizelge 3.29 646 kayıt için elde edilen en yüksek test doğruluk oranının (80,15%) konfüzyon matrisi.

		Gerçek Durum		
		Pozitif(1)	Negatif(0)	
Teşhis	n=196			
	Pozitif(1)	GP=3	YP=38	41
	Negatif(0)	YN=1	GN=154	155
		4	192	

Duyarlılık ve Özgüllük hesaplaması Eşitlik 3.1 Ve Eşitlik 3.2 'de gösterilmiştir.

$$\text{Duyarlılık} = \frac{GP}{GP+YN} \quad (3.1)$$

$$\text{Özgüllük} = \frac{GN}{GN+YP} \quad (3.2)$$

Eşitlik 3.1 ve 3.2'de yer alan GP; Gerçek Durum Pozitifi, GN; Gerçek Durum Negatifi, YP; Yanlış Durum Pozitifi, YN ise Yanlış Durum Negatifi ifade etmektedir. Duyarlılık bir test etme durumunda ya da teşhis yönteminde testin hastalığı taşıyanları hangi oran ile tespit edebildiğini belirtir. Özgüllük ise bir test etme durumunda ya da teşhis yönteminde hasta olmayan bir hastaya doğru teşhisi koyma oranını belirtir.

Çizelge 3.30 515 kayıt için elde edilen en yüksek test doğruluk oranının ( 82,40%) konfüzyon matrisi.

		Gerçek Durum		
		Pozitif(1)	Negatif(0)	
Teşhis	n=154			
	Pozitif(1)	GP=10	YP=22	32
	Negatif(0)	YN=5	GN=117	122
		15	139	

Çizelge 3.29 için Duyarlılık 75% ve Özgüllük 80,2%;

Çizelge 3.30 için Duyarlılık 66,6%ve Özgüllük 84,17% olarak hesaplanmıştır.

#### 4.SONUÇ

Bu tez çalışmasında Parkinson hastalığının tespiti için insan bilgisayar etkileşimi yaklaşımından hareket ederek tuş vuruş verileri kullanılmıştır. Günümüzde tuş vuruş dinamikleri; kimlik doğrulama sistemlerinden sağlık alanına, yaş tespitinden stres ve psikolojik analiz sentezine kadar pek çok alanda kullanılmaya başlanmıştır. Parkinson hastalığının kısa sürede ve kolay bir şekilde tespit edilmesi günümüzde ciddi bir problemdir. Yapılan bu çalışma ile kişilerin parkinson hastası olup olmadığını anlamak için maddi ve manevi zorluk yaratan testlerin önüne geçilmiştir. Bu çalışmadaki yöntem insan ile bilgisayar klavyesi arasındaki etkileşimden yola çıkılarak geliştirilmiştir. Geliştirilen yöntem zaman serisi analizi tekniklerini ve makine dili öğrenme algoritmalarını içermektedir. Uygulanan teknikler ile tez hedefleri yerine getirilmeye çalışmıştır.

Bu tez çalışmasında geliştirilen yöntemde ilk olarak tuş vuruş zaman serisinin doğrusal ve doğrusal olmayan dinamikleri incelenmiştir. Ardından, farklı özellikler ve farklı zaman serileri için elde edilen tüm sonuçlara istatistiksel analizler (ANOVA) uygulanmış ve ayırt edici olan özellikler belirlenmiştir.

Sınıflayıcıya sokulan özellikler eğitim ve test oranlarının farklı olduğu iki durum için denenmiştir. İlk durumda test oranı 50% ve eğitim oranı 50% olarak alınmıştır. İkinci durumda ise eğitim oranı 70% ve test oranı 30% alınarak alınmıştır.

İlk durumda ANOVA sonuçlarına göre ayırt edici olarak bulunan tüm özellikler teker teker iki sınıflayıcının girişine uygulandı. Bunun sonucunda elde edilen en yüksek test doğruluk oranları, Destek Vektör Makinesi sınıflayıcısında 646 kayıt için test oranı 79,04% ve eğitim oranı 77,55%, 515 kayıt için test oranı 78,50% ve eğitim oranı 77,06% iken; k En Yakın Komşu sınıflayıcısında ise 646 kayıt için test oranı 78,21% ve eğitim oranı 81,98%, 515 kayıt için test oranı 79,99% ve eğitim oranı 83,14% şeklindedir.

İkinci durumda; tüm ayırt edici çıkan özellikler içerisinde en ayırt edici olan 5 tanesi seçilerek beşerli şekilde iki sınıflayıcının girişine sokulmuştur. Bunun sonucunda elde edilen en yüksek doğruluk oranları, Destek Vektör Makinesi sınıflayıcısında 646 kayıt için test oranı 78,31% ve eğitim oranı 78,32%, 515 kayıt için test oranı 77,23% ve eğitim oranı 77,59% iken; k En Yakın Komşu sınıflayıcısında 646 kayıt için test oranı

76,14% ve eğitim oranı 79,84%, 515 kayıt için test oranı 80,15% ve eğitim oranı 83,24% şeklindedir.

Üçüncü durumda; ayırt edici olarak belirlenen tüm özellikler hepsi birlikte iki sınıflayıcının girişine sokulmuştur. Bunun sonucunda elde edilen en yüksek doğruluk oranları, Destek Vektör Makinesi sınıflayıcısında 646 kayıt için test oranı 78,19% ve eğitim oranı 78,38%, 515 kayıt için test oranı 76,86% ve eğitim oranı 78,48% iken; k En Yakın Komşu sınıflayıcısında 646 kayıt için test oranı 79,58% ve eğitim oranı 81,96%; 515 kayıt için test oranı 80,12% ve eğitim oranı ise 83,07% şeklindedir.

Dördüncü ve son durumda ise; ayırt edici olarak belirlenen özellikler içerisinde P değeri 0,001'e eşit veya küçük olanlar seçilmiş ve bunlara Rasgele Orman Algoritması uygulanmıştır. Bu algoritma ile birlikte bu özellikler içerisinde en değersiz olanları belirlenerek her defasında tüm özellikler içerisinde çıkartılmış ve sınıflayıcı doğruluk oranları hesaplanmıştır. Bunun sonucunda en değerli özellik 646 ve 515 kayıt için Basılı Tutma Süresi LR-RL Asimetri özelliği ortak çıkarken, en değersiz özellik 646 kayıt için Geçiş Süresi durum uzayı IV. bölge nokta sayısı, 515 kayıt için Gecikme Süresi Histogram 2. Alan özelliği çıkmıştır. Destek Vektör Makinesi'ne Rasgele Orman Algoritması uygulandığı durumda elde edilen en yüksek doğruluk oranı 646 kayıt için test oranı 77,52% ve eğitim oranı 79,13%, 515 kayıt için test oranı 78,05% ve eğitim oranı 77,17 iken k En Yakın Komşu sınıflayıcısına uygulandığı zaman 646 kayıt için test oranı 80,15% ve eğitim oranı 83,78%, 515 kayıt için ise test oranı 82,4% ve eğitim oranı 86,64% şeklindedir.

Tuş vuruş datalarını kullanarak Parkinson hastalığını tespit etmeye çalışan diğer çalışmalar incelendiğinde; L. Giancardo ve arkadaşları 81%, Warwick R. Adams ve arkadaşları ise 98% doğruluk oranı elde etmişlerdir. Warwick R. Adams ve arkadaşlarının yaptıkları çalışmada doğruluk oranının, yapılan bu çalışmadan daha yüksek çıkmasının sebebi; çalışma süresince hangi tuşlara basıldığına bilinmesidir. Yapılan bu çalışmada katılımcıların hangi tuşlara bastığı bilinmemekle birlikte, kayıt sırasında meydana gelen duraksamaların geçiş sürelerini etkilediği düşünülmektedir. Elde edilen sonuçlar detaylı olarak incelendiğinde yapılan bu çalışmanın diğer sensör ölçümleri ile Parkinson hastalığını tespit etmeye çalışan çalışmalar yerine kullanılabileceği ispatlanmıştır.

Bu alıřmanın sonunda elde edilen doęruluk oranını daha da artırılabilir. Mevcut alıřmanın doęruluk oranını ve güvenilirliğini artırmak iin eęitim setini artırmak bir özüm olabilir. Bu sayıyı artırmak iin gönüllü katılımcılar gerekmektedir. Veri setindeki zorluklar göz önüne alındığında sadece basılı tutma süresine odaklanarak doęruluk oranı daha da artırılabilir. Bir dięer yöntem ise yeni ayırt edici özellik bulmaktır. Sinyal işleme teknikleriyle bu alıřmada daha önce incelenmeyen; özellikle doğrusal olmayan dinamikler detaylı şekilde incelenmelidir. Mevcut alıřma farklı sınıflayıcılar ile ve farklı eęitim / test oranı ile de denenebilir.

## 5. KAYNAKLAR LİSTESİ

- [1] Ozen, Zeki. (2016). Kimlik Doğrulaması için Tuş Vuruş Dinamiklerine Dayalı Bir Güvenlik Sisteminin Yapay Sinir Ağları ile Geliştirilmesi. 10.13140/RG.2.2.34657.10082.
- [2] Pentel, Avar. (2017). Emotions and User Interactions with Keyboard and Mouse. 10.1109/IISA.2017.8316379.
- [3] STRESS DETECTION FOR KEYSTROKE DYNAMICS SHING-HON LAU May 2018 CMU-ML-18-104 Machine Learning Department School of Computer Science Carnegie Mellon University Pittsburgh, PA
- [4] Pentel, Avar. (2018). Predicting User Age by Keystroke Dynamics.
- [5] P. Grosse, M.J. Cassidy, P. Brown, EEG–EMG, MEG–EMG and EMG–EMG frequency analysis: Physiological principles and clinical applications, Clinical Neurophysiology, Volume 113, Issue 10, 2002, Pages 1523-1531, ISSN 1388-2457, [https://doi.org/10.1016/S1388-2457\(02\)00223-7](https://doi.org/10.1016/S1388-2457(02)00223-7). (<http://www.sciencedirect.com/science/article/pii/S1388245702002237>)
- [6] Unified Parkinson's Disease Rating Scale (UPDRS) and The Movement-Disorder Society Sponsored –Unified Parkinson's Disease Rating Scale (MDS-UPDRS)
- [7] Teaching program for the Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale: (MDS-UPDRS).  
Goetz CG<sup>1</sup>, Stebbins GT, Chmura TA, Fahn S, Poewe W, Tanner CM.
- [8] Automatic detection of freezing of gait events in patients with Parkinson's disease. Evanthia E. Tripoliti a, Alexandros T. Tzallas a, Markos G. Tsipouras a, George Rigas a, Panagiota Bougiaa, Michael Leontioub, Spiros Konitsiotisc, Maria Chondrogiorgi c, Sofia Tsouli c, Dimitrios I. Fotiadis DI.
- [9] Mazilu, Sînziana & Blanke, Ulf & Hardegger, Michael & Troster, Gerhard & Gazit, Eran & Dorfman, Moran & Hausdorff, Jeffrey. (2014). GaitAssist: A wearable assistant for gait training and rehabilitation in Parkinson's disease. 2014 IEEE International Conference on Pervasive Computing and Communication Workshops, PERCOM WORKSHOPS 2014. 135-137. 10.1109/PerComW.2014.6815179.
- [10] Detecting freezing of gait with a tri-axial accelerometer in Parkinson's disease patients. Claas Ahlrichs<sup>1</sup> · Albert Samà<sup>3</sup> · Michael Lawo<sup>2</sup> · Joan Cabestany<sup>3</sup> · Rodríguez-Martín<sup>3</sup> · Carlos Pérez-López<sup>3</sup> · Dean Sweeney<sup>4</sup> · Leo R. Quinlan<sup>4</sup> · Gearòid Ó Laighin<sup>4</sup> · Timothy Counihan<sup>5</sup> · Patrick Browne<sup>5</sup> · Lewy Hadas<sup>6</sup> · Gabriel Vainstein<sup>6</sup> · Alberto Costa<sup>7,9</sup> · Roberta Annicchiarico<sup>7</sup> · Sheila Alcaine<sup>8</sup> · Berta Mestre<sup>8</sup> · Paola Quispe<sup>8</sup> · Àngels Bayes<sup>8</sup> · Alejandro Rodríguez-Molinero<sup>4</sup>

- [11] Sang-Hong Lee and Joon S. Lim. 2012. Parkinson's disease classification using gait characteristics and wavelet-based feature extraction. *Expert Syst. Appl.* 39, 8 (June 2012), 7338-7344. DOI=<http://dx.doi.org/10.1016/j.eswa.2012.01.084>
- [12] Quantitative estimation of foot-flat and stance phase of gait using foot-worn inertial sensors. Benoit Mariani<sup>a</sup> Hossein Rouhani<sup>a</sup> Xavier Crevoisier<sup>b</sup> Kamiar Aminian<sup>a</sup>
- [13] AN AMBULATORY SYSTEM TO QUANTIFY BRADYKINESIA AND TREMOR IN PARKINSON'S DISEASE A. Salarian<sup>1</sup>, H. Russmann<sup>2</sup>, F. J. G. Vingerhoets<sup>2</sup>, P. R. Burkhard<sup>3</sup>, Y. Blanc<sup>3</sup>, C. Dehollain<sup>4</sup>, Member, IEEE, K. Aminian<sup>1</sup>, Member, IEEE
- [14] Advanced Analysis of Finger-Tapping Performance: A Preliminary Study. Çağatay Barut,<sup>1</sup> Erhan Kızıltan,<sup>2</sup> Ethem Gelir,<sup>3</sup> and Füzün Köktürk<sup>4</sup>
- [15] Carpal tunnel syndrome and the use of computer mouse and keyboard A Review Jane Frølund Thomsen Arbejdsmedicinsk Klinik, Amtssygehuset i Glostrup
- [16] Towards a User-Friendly Brain-Computer Interface: Initial Tests in ALS and PLS Patients Ou Bai,<sup>1,3,\*</sup> Peter Lin,<sup>2</sup> Dandan Huang,<sup>1</sup> Ding-Yu Fei,<sup>1</sup> and Mary Kay Floeter<sup>4</sup>
- [17] Computer keyboard interaction as an indicator of early Parkinson's disease. L. Giancardo<sup>1,\*</sup>, A. Sánchez-Ferro<sup>1,2,3,4,5,\*</sup>, T. Arroyo-Gallego<sup>1,6</sup>, I. Butterworth<sup>1</sup>, C. S. Mendoza<sup>1</sup>, P. Montero<sup>7</sup>, M. Matarazzo<sup>2,3,4,5</sup>, J.A. Obeso<sup>2,3,4</sup>, M. L. Gray<sup>1,8</sup> & R. San José Estépar<sup>9</sup>
- [18] High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing Warwick R. Adams.
- [19] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and Physionet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220 [Circulation Electronic Pages; <http://circ.ahajournals.org/content/101/23/e215>]; 2000 (June 13).)
- [20] Dawson, B; Trapp, R (2001). *Basic & Clinical Biostatistics* (4th Edition ed.). McGrawHill.
- [21] What is Kurtosis? An Influence Function Approach David Ruppert a a University of North Carolina, Chapel Hill, NC, 27514, USA
- [22] An Influence Function Approach to Describing the Skewness of a Distribution. Richard A. Groeneveld
- [23] Entropy – A Guide for the Perplexed Roman Frigg and Charlotte Werndl\* June 2010
- [24] A Mathematical Theory of Communication\*. C. E. Shannon.

- [25] Approximate entropy (ApEn) as a complexity measure. Steve Pincus
- [26] Entropy and Mutual Information Erik G. Learned-Miller Department of Computer Science University of Massachusetts, Amherst Amherst, MA 01003 September 16, 2013
- [27] Information Theory, Inference, and Learning Algorithms David J.C. MacKay
- [28] Independent Coordinates For Strange Attractors From Mutual Information. Andrew M. Fraser and Harry L. Swinney. Department of Physics, University of Texas, Austin, Texas. 1985
- [29] Kennel, M.; Brown, R.; Abarbanel, H. (1992). "Determining embedding dimension for phase-space reconstruction using a geometrical construction". *Physical Review A*. 45 (6): 3403–3411. Bibcode:1992PhRvA..45.3403K. doi:10.1103/PhysRevA.45.3403. PMID 9907388.
- [30] Rhodes, C.; Morari, M. (1997). "The false nearest neighbors algorithm: An overview". *Computers & Chemical Engineering*. 21: S1149. doi:10.1016/S0098-1354(97)87657-0.
- [31] Guyon, Isabelle & Elisseeff, André. (2003). An Introduction of Variable and Feature Selection. *J. Machine Learning Research Special Issue on Variable and Feature Selection*. 3. 1157 - 1182. 10.1162/153244303322753616.
- [32] Lindman, H. R. (1974). *Analysis of variance in complex experimental designs*. Oxford, England: W. H. Freeman & Co.
- [33] Tin Kam Ho. 1995. Random decision forests. In *Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1) - Volume 1 (ICDAR '95)*, Vol. 1. IEEE Computer Society, Washington, DC, USA, 278-.
- [34] Support-Vector Networks. Corina Cortes, Vladimir Vapnik
- [35] The Comparison of the ANN and SVM Methods for the Prediction of Voltage Drop on a Subway Line. İlhan Kocaarslan, Mehmet Taciddin Akçay, Abdurrahim Akgündoğdu, Hasan Tiryaki. İstanbul 2017
- [36] Cristianini, N., and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, First Edition (Cambridge: Cambridge University Press).
- [37] Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi. Sevgi Ayhan, Şenol Erdoğmuş. 2014
- [38] Altman, N.S.. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *American Statistician - AMER STATIST*. 46. 175-185. 10.1080/00031305.1992.10475879.
- [39] Introduction to machine learning: k-nearest neighbors Zhongheng Zhang

- [40] Soman, K.p., Loganathan, R. and Ajay, V. (2011). Machine learning with SVM and other kernel methods. PHI Learning Pvt. Ltd., 286 s.
- [41] Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition, data mining and knowledge discovery. Kluwer Academic Publishers, 2 (2), 121-167
- [42] Gunn, S. R. (1998). Support vector machines for classification and regression. Technical Report, Faculty of Engineering, Science and Mathematics, School of Electronics and Computer Science. <http://users.ecs.soton.ac.uk/srg/publications/pdf/SVM.pdf>
- [43] Cortes, C., Vapnik, V. (1995). Support vector networks, Machine Learning, 20,1-25.
- [44] Busuttill, S. (2003). Support vector machines. In Proceedings of the Computer Science Annual Research Workshop, Villa Bighi, Kalkara, University of Malta
- [45] Suykens, J. A. K. (2002). Least squares support vector machines. River Edge, NJ: World Scientific, xiv, 294 s.