

T.C.
ADNAN MENDERES ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK YÜKSEK LİSANS PROGRAMI

**EEG SİNYALLERİNDE FARKLI BOYUT İNDİRGEME VE
SINIFLANDIRMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI**

HAKAN ÖZTÜRK
YÜKSEK LİSANS TEZİ

DANIŞMAN
Prof. Dr. Mevlüt TÜRE

II. DANIŞMAN
Prof. Dr. Nefati KIYLIOĞLU

AYDIN-2016

KABUL VE ONAY SAYFASI

T.C. Adnan Menderes Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı Yüksek Lisans Programı çerçevesinde Hakan ÖZTÜRK tarafından hazırlanan “EEG Sinyallerinde Farklı Boyut İndirgeme ve Sınıflandırma Yöntemlerinin Karşılaştırılması” başlıklı tez, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi:/...../.....

Ünvanı, Adı ve Soyadı	Üniversitesi	İmza
Prof. Dr. Mevlüt TÜRE (Tez Danışmanı)	Adnan Menderes Üniversitesi	
Prof. Dr. Mehmet N. ORMAN	Ege Üniversitesi	
Doç. Dr. İmran KURT ÖMÜRLÜ	Adnan Menderes Üniversitesi	

ONAY:

Bu tez Adnan Menderes Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca yukarıdaki jüri tarafından uygun görülmüş ve Sağlık Bilimleri Enstitüsününtarih vesayılı oturumunda alınannolu Yönetim Kurulu kararıyla kabul edilmiştir.

Prof. Dr. Ahmet CEYLAN
Enstitü Müdürü

TEŞEKKÜR

Yüksek lisans öğrenimim süresince desteğini hiçbir zaman esirgemeyen, bilgi ve birikimiyle tez çalışmamın tüm aşamalarında yanımda olan ve bana yol gösteren tez danışmanım Prof. Dr. Mevlüt TÜRE'ye,

Tez çalışmam süresince desteğini esirgemeyen, bilgi ve birikimiyle tezime katkıda bulunan ikinci tez danışmanın Prof. Dr. Nefati KIYLIOĞLU'na,

Yüksek lisans öğrenimim süresince yardımını esirgemeyen ve mesleki gelişimime katkıda bulunan Doç. Dr. İmran KURT ÖMÜRLÜ'ye,

Çalışma arkadaşım Fulden CANTAŞ ve kuzenim Özge ULUSKARADAĞ'a,

Desteklerini hiçbir zaman esirgemeyen varlıklarıyla bana güç veren annem Gülhan ÖZTÜRK ve babam Mustafa ÖZTÜRK'e teşekkür ederim.

İÇİNDEKİLER

	Sayfa
KABUL VE ONAY SAYFASI.....	i
TEŞEKKÜR	ii
İÇİNDEKİLER.....	iii
SİMGELER ve KISALTMALAR DİZİNİ.....	vi
ŞEKİLLER DİZİNİ	viii
TABLolar DİZİNİ.....	xi
ÖZET	xii
ABSTRACT	xiii
1. GİRİŞ.....	1
2. GENEL BİLGİLER.....	5
2.1 Beyin.....	5
2.2 Elektroansefalogram (EEG)	7
2.2.1 EEG'nin Tarihçesi	7
2.2.2 EEG'nin Tanımı	8
2.2.3 EEG Sinyallerinin Oluşumu ve Ölçümü	8
2.3 Epilepsi	11
2.4 Epilepsinin Bilgisayar Destekli Teşhisi.....	14
2.5 Sinyal İşleme Yöntemleri	15
2.5.1 Spektral Analiz	15
2.5.1.1 Periyodisite	16
2.5.1.2 Fourier Serisi ve Fourier Dönüşümü	18
2.5.1.3 Kısa Zamanlı Fourier Dönüşümü	20
2.5.1.4 Dalgacık Dönüşümü	21
2.5.1.4.1 Sürekli Dalgacık Dönüşümü.....	22

2.5.1.4.2 Ayrık Dalgacık Dönüşümü.....	25
2.6 Boyut İndirgeme Yöntemleri.....	29
2.6.1 Temel Bileşenler Analizi	29
2.6.2 Bağımsız Bileşen Analizi	31
2.7 Sınıflandırma Yöntemleri.....	35
2.7.1 Diskriminant Analizi	35
2.7.1.1 Diskriminant Analizine Geometrik Yaklaşım	36
2.7.1.2 İki Grup için Lineer Diskriminant Analizi	37
2.7.1.3 Lineer Diskriminant Analizinde Gözlemlerin Sınıflandırılması	40
2.7.2 Destek Vektör Makinesi	43
2.7.2.1 Lineer Olarak Ayrılabilir Durumu için DVM.....	45
2.7.2.2 Lineer Olarak Belirli Bir Hata ile Ayrılabilir Durumu için DVM.....	48
2.7.2.3 Lineer Olmayan DVM.....	50
2.7.3 Sınıflandırma Başarısı Ölçütleri	51
3. GEREÇ VE YÖNTEM.....	53
3.1 Verilerin Toplanması.....	53
3.2 Veri Setlerinin Oluşturulması.....	54
3.3 Özniteliklerin Çıkarılması ve Boyutunun İndirgenmesi.....	57
4. BULGULAR	64
4.1 Özniteliklerin TBA ve BBA ile Boyutlarının İndirgenmesi.....	64
4.2 LDA Sınıflandırma Sonuçları.....	66
4.3 DVM Sınıflandırma Sonuçları.....	72
4.3.1 Lineer-DVM Sınıflandırma Sonuçları.....	72
4.3.2 Radyal-DVM Sınıflandırma Sonuçları.....	77
5. TARTIŞMA.....	83
6. SONUÇ ve ÖNERİLER	87
7. KAYNAKLAR.....	89

8. EKLER	95
Ek 1	95
ÖZGEÇMİŞ	



SİMGELER ve KISALTMALAR DİZİNİ

ADD	: Ayrık Dalgacık Dönüşümü
BBA	: Bağımsız Bileşenler Analizi
BYSM	: Beklenen Yanlış Sınıflandırma Maliyeti
DA	: Diskriminant Analizi
DD	: Dalgacık Dönüşümü
DVM	: Destek Vektör Makinesi
EEG	: Elektroensefelografi
FD	: Fourier Dönüşümü
Hz	: Hertz
KZFD	: Kısa Zamanlı Fourier Dönüşümü
LDA	: Lineer Diskriminant Analizi
SDD	: Sürekli Dalgacık Dönüşümü
TBA	: Temel Bileşenler Analizi
$\psi(t)$	: Ana Dalgacık Fonksiyonu
τ	: Öteleme Parametresi
μV	: Mikrovolt
$X(f)$	: Fourier Dönüşümü
$W(s, \tau)$	: Dalgacık Katsayıları
Var	: Varyans
C	: Kabul Edilebilirlik Sabiti
$\mathbf{a}$	: Ağırlık Vektörü
$\mathbf{I}$	: Birim Matris
ϕ	: Faz
θ	: Teta
δ	: Delta
γ	: Genişleme Parametresi
β	: Beta
α	: Alfa
Z	: Diskriminant Fonksiyonu
w	: Hiper Düzlem
$\mathbf{W}$	: Ayırıştırma Matrisi

t	: Zaman
T	: Periyot
s	: Ölçek Parametresi
K	: Çekirdek Fonksiyonu
J	: Youden İndeksi
h	: Alçak Geçiren Filtre
g	: Yüksek Geçiren Filtre
f	: Frekans
E	: Beklenen Değer
$db4$	: Daubechies 4 Dalgacığı
D	: Detay Katsayıları
$c(i j)$	: Yanlış Sınıflandırma Maliyeti
C	: Düzenleyici Parametre
A	: Yaklaşım Katsayıları
$\mathbf{A}$	: Karıştırma Matrisi
$\varphi_{m,k}(n)$	: Ölçek Fonksiyonu
ξ_i	: Sınıflandırma Hataları
λ_i	: i. Öz Değer
α_i	: Lagrange Çarpanları
$\bar{x}$	: Ortalama
$\mathbf{v}_i$	: i. Öz Vektör
$\mathbf{m}_x$	: Ortalama Vektörü
$\mathbf{S}_{\text{ortak}}$	: Ortak Varyans Kovaryans Matrisi
$\mathbf{R}_x$	: Korelasyon Matrisi
$\mathbf{C}_x$	: Kovaryans Matrisi

ŞEKİLLER DİZİNİ

	Sayfa
Şekil 2.1 Beyin lobları ve beyincik	6
Şekil 2.2 Kaydedilen ilk EEG.....	7
Şekil 2.3 Sinir hücresinin genel yapısı	9
Şekil 2.4 Aksiyon potansiyeli.....	10
Şekil 2.5 Standart 10-20 elektrot yerleşimi	11
Şekil 2.6 EEG sinyallerinin beta, alfa, teta ve delta bantları	13
Şekil 2.7 Bilgisayar destekli karar sistemlerinde süreç	14
Şekil 2.8 Periyodik bir fonksiyonun grafiği ve parametreleri	16
Şekil 2.9 Durağan olmayan bir zaman serisinin alabileceği durumlar	17
Şekil 2.10 Fourier Dönüşümü.....	19
Şekil 2.11 Sinyalin pencerelenmesi.....	20
Şekil 2.12 Pencerelenen sinyale uygulanan Fourier Dönüşümü	21
Şekil 2.13 Düşük ve yüksek ölçek.....	25
Şekil 2.14 Sinyalin ayrıştırılması.....	26
Şekil 2.15 ADD ile bir sinyalin alt bantlarına ayrıştırılması	27
Şekil 2.16 Sinyalin beşinci seviye ADD yöntemiyle alt bantlarına ayrıştırılması	28
Şekil 2.17 Daubechies Dalgacıklarından db4 Dalgacığı	29
Şekil 2.18 BBA modeli.....	31
Şekil 2.19 Beyazlatma ve BBA dönüşümü	34
Şekil 2.20 İki gruplu diskriminant analizinin grafik gösterimi	36
Şekil 2.21 Ayırıcı hiper düzlemler	43
Şekil 2.22 Destek vektörleri ve optimum hiper düzlem	44
Şekil 2.23 Lineer ayrılabilme durumu.....	46
Şekil 2.24 Doğrusal olarak belirli bir hata ile ayrılabilen veri kümesi için DVM sınıflandırıcısı.....	48
Şekil 2.25 Giriş uzayının doğrusal sınıflamanın mümkün olduğu çok boyutlu uzaya dönüşümü	50
Şekil 3.1 Epilepsi tanısı konmuş bir hastaya ait EEG sinyali.....	53
Şekil 3.2 Sağlam bir bireye ait EEG sinyali	54
Şekil 3.3 EEG kayıtlarının birleştirilmesi.....	54
Şekil 3.4 Hasta bireylere ait 300 sn uzunluğundaki birleştirilmiş EEG sinyalleri	55

Şekil 3.5 Sağlam bireylere ait 300 sn uzunluğundaki birleştirilmiş EEG sinyalleri	56
Şekil 3.6 $x(n)$ sinyaline uygulanan 6. seviye ADD sonucu elde edilen alt bantlar ve bu alt bantlara ilişkin frekans aralıkları.....	58
Şekil 3.7 Epileptik sinyalin 6. seviye ADD yöntemiyle alt bantlarına ayrıştırılması ile elde edilen yaklaşım ve detay alt bantları.....	59
Şekil 3.8 Bir saniyelik çerçeveden elde edilen D4, D5, D6 ve A6 alt bantları ve bu alt bantlardan hesaplanan istatistikler	61
Şekil 3.9 Öznitelik matrisi	62
Şekil 3.10 İşlem sürecinin akış diyagramı.....	63
Şekil 4.1 Özniteliklerin diskriminant skorları histogram grafiği.....	66
Şekil 4.2 Özniteliklerin diskriminant skorları histogram grafiği.....	67
Şekil 4.3 Temel bileşenlerin diskriminant skorları histogram grafiği	68
Şekil 4.4 Temel bileşenlerin diskriminant skorları histogram grafiği	68
Şekil 4.5 Bağımsız bileşenlerin diskriminant skorları histogram grafiği	69
Şekil 4.6 Bağımsız bileşenlerin diskriminant skorları histogram grafiği	70
Şekil 4.7 Eğitim veri setlerinin LDA ile yapılan sınıflandırma sonuçlarının karşılaştırılması	71
Şekil 4.8 Test veri setlerinin LDA ile yapılan sınıflandırma sonuçlarının karşılaştırılması	71
Şekil 4.9 Özniteliklerle yapılan sınıflandırmada eğitim ve test veri setleri için özniteliklerin sınıf sınırlarından uzaklıkları	73
Şekil 4.10 Temel bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için temel bileşenlerin sınıf sınırlarından uzaklıkları.....	74
Şekil 4.11 Bağımsız bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için bağımsız bileşenlerin sınıf sınırlarından uzaklıkları	75
Şekil 4.12 Eğitim veri setlerinin Lineer-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırılması	76
Şekil 4.13 Test veri setlerinin Lineer-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırılması	76
Şekil 4.14 Özniteliklerle yapılan sınıflandırmada eğitim ve test veri setleri için özniteliklerin sınıf sınırlarından uzaklıkları	77
Şekil 4.15 Temel bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için temel bileşenlerin sınıf sınırlarından uzaklıkları.....	78
Şekil 4.16 Bağımsız bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için gözlemlerin sınıf sınırlarından uzaklıkları	79

Şekil 4.17 Eğitim veri setlerinin Radyal-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırması	80
Şekil 4.18 Test veri setlerinin Radyal-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırması	81
Şekil 4.19 Özniteliklerin sınıflandırma yöntemlerine göre karşılaştırma grafiği	81
Şekil 4.20 Temel bileşenlerin sınıflandırma yöntemlerine göre karşılaştırma grafiği	82
Şekil 4.21 Bağımsız bileşenlerin sınıflandırma yöntemlerine göre karşılaştırma grafiği	82



TABLÖLAR DİZİNİ

	Sayfa
Tablo 2.1 EEG sinyallerinin frekans bantları	12
Tablo 2.2 Yanlış sınıflandırma maliyeti	41
Tablo 2.3 Sınıflandırma tablosu	51
Tablo 3.1 Altıncı seviye dalgacık ayrışımındaki frekans bantlarının aralıkları	60
Tablo 4.1 Temel bileşenlerin öz değerleri, varyans ve kümülatif varyans yüzdeleri	64
Tablo 4.2 Bağımsız bileşenlerin öz değerleri, varyans ve kümülatif varyans yüzdeleri	65
Tablo 4.3 Özniteliklere ait eğitim ve test veri setlerinin LDA sınıflandırma sonuçları	67
Tablo 4.4 Temel bileşenlere ait eğitim ve test veri setlerinin LDA sınıflandırma sonuçları ...	69
Tablo 4.5 Bağımsız bileşenlere ait eğitim ve test veri setlerinin LDA sınıflandırma sonuçları	70
Tablo 4.6 Özniteliklere ait eğitim ve test veri setlerinin Lineer-DVM sınıflandırma sonuçları	73
Tablo 4.7 Temel bileşenlere ait eğitim ve test veri setlerinin Lineer-DVM sınıflandırma sonuçları	74
Tablo 4.8 Bağımsız bileşenlere ait Eğitim ve test veri setlerinin Lineer-DVM sınıflandırma sonuçları	75
Tablo 4.9 Özniteliklere ait eğitim ve test veri setlerinin Radyal-DVM sınıflandırma sonuçları	78
Tablo 4.10 Temel bileşenlere ait Eğitim ve test veri setlerinin Radyal-DVM sınıflandırma sonuçları	79
Tablo 4.11 Bağımsız bileşenlere ait eğitim ve test veri setlerinin Radyal-DVM sınıflandırma sonuçları	80

ÖZET

EEG SİNYALLERİNDE FARKLI BOYUT İNDİRGEME VE SINIFLANDIRMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Öztürk H. Adnan Menderes Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı Yüksek Lisans Tezi, Aydın, 2016.

Bu çalışmada, epileptik ve epileptik olmayan EEG sinyallerinden elde edilen özniteliklerin boyutlarının Temel Bileşenler Analizi (TBA) ve Bağımsız Bileşenler Analizi (BBA) yöntemleri ile indirgenmesinin sınıflandırma başarısı üzerine etkilerinin belirlenmesi ve Lineer Diskriminant Analizi (LDA) ile Destek Vektör Makinesi (DVM) yöntemlerinin sınıflandırma performanslarının karşılaştırılması hedeflendi.

Çalışmaya 10 kontrol ve uzman hekim tarafından epilepsi tanısı konmuş 10 hasta olmak üzere toplam 20 kişi dahil edildi. Epilepsi tanısı konmuş hastalardan alınan EEG kayıtları nöbet geçirmediikleri sırada alınan kayıtlardı. Epileptik ve epileptik olmayan sinyalleri sınıflandırmak için Ayrık Dalgacık Dönüşümü (ADD) ile sinyallerinin spektral analizi gerçekleştirildi ve sınıflandırmada kullanılacak olan öznitelikler elde edildi. Öncelikle özniteliklerin boyutu indirgenmeden, daha sonra TBA ve BBA ile indirgenerek sınıflandırma yapıldı. Sınıflandırma LDA, lineer ve radyal tabanlı çekirdek fonksiyonlarının kullanıldığı DVM yöntemleri ile gerçekleştirildi.

Sonuç olarak, en yüksek sınıflandırma başarısı %92,2 duyarlılık %85,6 özgüllük ve %88,9 doğruluk oranlarıyla özniteliklerde boyut indirgenme yapılmadan ve radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM yöntemi ile elde edildi. BBA ve TBA ile boyutu indirgenen özniteliklerle yapılan sınıflandırmalarda benzer sonuçlar elde edildi.

Anahtar Kelimeler: Temel Bileşenler Analizi, Bağımsız Bileşen Analizi, Ayrık Dalgacık Dönüşümü, Destek Vektör Makinesi, Lineer Diskriminant Analizi, Sınıflandırma

ABSTRACT

THE COMPARISON OF VARIOUS DIMENSION REDUCTION AND CLASSIFICATION METHODS IN EEG SIGNALS

**Öztürk H. Adnan Menderes University Institute of Medical Sciences Biostatistics
Department Master of Science Thesis, Aydın, 2016.**

In this study, it is aimed to find out the impact of features derived from epileptic and non epileptic EEG signals and the reduction of their dimensions by Principal Component Analysis (PCA) and Independent Component Analysis (ICA) on classification success and to compare the classification performance of Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) are used.

A total of 20 people, 10 epilepsy patients who were diagnosed by specialist physicians and 10 non-epilepsy patients were included in the study. The EEG recordings taken from those who were diagnosed with epilepsy while they were not having seizures. To classify epileptic and non epileptic signals, first and foremost Discrete Wavelet Transform (DWT) and the spectral analysis of EEG signals were performed and features that would be used to classify the signals were obtained. First, features were classified without their dimensions being reduced. Then, they were classified after their dimensions were reduced by PCA and ICA. The classification was performed by LDA and SVM in which linear and radial basis kernel functions are used.

As a result, the highest classification success with 92.5% sensitivity, 85.6% specificity, and 88.9% accuracy ratios were obtained by DVM method in which features, whose dimensions were not reduced and radial basis kernel function are used. Similar classification results were obtained by features whose dimensions were reduced by PCA and ICA.

Keywords: Principal Components Analysis, Independent Component Analysis, Discrete Wavelet Transform, Support Vector Machine, Linear Discriminant Analysis, Classification

1. GİRİŞ

Beyin, insan vücudunun en karmaşık yapıya sahip olan organıdır. İnsan beyinde yaklaşık olarak 100 milyar beyin hücresi (nöron) bulunmaktadır. Hareket etme, konuşma, nefes alma, duyu organlarıyla hissetme gibi motor aktiviteleri ve bilişsel faaliyetleri kontrol eden beyin tarafından, bütün vücuda bu işlevleri gerçekleştirebilmek için farklı frekanslarda ve çok düşük miktarda elektriksel sinyaller salınmaktadır (Nunez ve Srinivasan, 2006). Beynin elektriksel aktivitelerinde meydana gelen bazı değişiklikler ve düzensizlikler bir takım hastalıkların habercisi olabilmektedir. Bu hastalıklardan bir tanesi de halk arasında sara hastalığı olarak bilinen epilepsidir. Epilepsi, beyinde bulunan sinir hücrelerinde nöbet (atak) adı verilen olağan dışı ve geçici elektrik boşalmaları sonucu meydana gelen, vücutta kontrol edilemeyen titreme, morarma, nefes darlığı, bayılma gibi istem dışı hareketlere yol açan nörolojik bir hastalıktır (Ropper ve Brown, 2005; Ropper, 2009). Dünya Sağlık Örgütü (DSÖ), Mayıs 2015'te dünya çapında yaklaşık olarak 50 milyon epilepsi hastası olduğunu ve bu nedenle epilepsinin en önemli nörolojik hastalıklardan birisi olduğunu ifade etmiştir.

Elektroensefelogram (EEG) beyinde oluşan elektriksel akımların Elektroensefelograf cihazı ile ölçülerek sayısallaştırılmış biçimidir. Özellikle epilepsi hastalığına tanı konulmasında ve tiplerinin belirlenmesinde EEG sinyallerinden yararlanılmaktadır (Nunez ve Srinivasan, 2006; Ropper, 2009). Epilepsi hastalığı, sebebi tam olarak anlaşılamayan bir biçimde oluşan nöbetlerle kendini gösterir. Ancak kısa süren epileptik dalga şekillerinin gözle tanınması zor ve zaman alıcı bir işlemdir (Mohammad ve ark, 2002). EEG kayıtlarının dikkatli bir şekilde analiz edilmesi, yaygın olan epilepsi hastalığının tanı ve tedavisi için oldukça önemlidir. Bu nedenle EEG cihazı nöroloji kliniklerinde beyin ile ilgili rahatsızlıkların tespit edilmesinde sıklıkla kullanılmaktadır (Sanei ve Chambers, 2007). Ancak uzmanların EEG kayıtlarından elde ettikleri bulguları analiz etme ve sınıflandırma olanakları kısıtlıdır. Uzmanların yaptığı değerlendirmelerde net ölçütlerin olmaması EEG sinyallerinin görsel analizinin yetersiz kalmasına neden olmaktadır. Rutin klinik teşhislerde EEG sinyallerinin analizine duyulan gereksinim, otomasyon ve bilgisayar tekniklerinin kullanımını gerekli kılmıştır. Bu sebeple nesnel bir değerlendirmede bulunabilmek için EEG sinyallerinin farklı tekniklerle analiz edilmesi yoluna gidilmiştir (Bronzino, 2000; Subaşı ve Gürsoy, 2010).

EEG sinyallerinde beynin fonksiyonları ile ilgili çok miktarda bilgi saklıdır. Bu bilgilerin kullanılabilmesi için sinyal işleme yöntemleri kullanılarak çeşitli araştırmalar yapılmaktadır. Yapılan araştırmalar, EEG sinyallerinin frekansının kişinin zihinsel aktivitelerine göre çeşitlilik gösterdiğini ortaya koymuştur (Liu ve ark, 2011, Chen, 2014).

EEG sinyallerinin en karakteristik özelliği durağan ve doğrusal olmayan ani değişimler göstermesidir. EEG sinyallerinde, genlik, faz ve frekans değerleri sürekli değişir. Herhangi bir EEG kaydı göz ile incelendiğinde, bireyin hasta olup olmadığını söylemek kolay değildir. Bu sebeple birçok çalışmada sağlıklı ve hasta bireylerin EEG verileri farklı istatistiksel yöntemlerle işlenerek analiz edilmiş ve belirgin farklar elde edilmeye çalışılmıştır (Yuan ve ark, 2008; Lima ve ark, 2010; Subaşı ve Gürsoy, 2010; Wang ve ark, 2013).

Literatürde yer alan birçok çalışmada, EEG sinyallerinin spektral analizinde Fourier Dönüşümü (FD) kullanılmıştır (Kıymık ve ark, 2004; Andrzejak ve ark, 2011; Coşkun ve İstanbullu, 2012). FD ile zaman bölgesinde bulunan bilgi, frekans bölgesine taşınmaktadır. Ancak FD, sinyallerin periyodik bir yapıya sahip olduğunu ve analiz boyunca karakteristiğinin değişmediğini kabul etmektedir. Bundan ötürü durağan sinyaller için başarılı sonuçlar elde edilse de durağan olmayan sinyaller için uygun bir analiz yöntemi değildir (Merry, 2005). Bu nedenle EEG sinyallerinin zamana bağlı olarak nasıl değiştiğini anlayabilmek ve bu sinyallerden anlamlı bir veri elde edebilmek amacıyla tüm sinyal bileşenlerini inceleyebilecek, durağan olmayan sinyal analizi yöntemlerine ihtiyaç duyulmaktadır. Biyolojik sinyallerin incelenmesinde Dalgacık Dönüşümü (DD) FD'nin eksikliklerini gidermek için geliştirilmiş bir yöntemdir. Bu analiz yöntemi gürültüye karşı daha az duyarlılık göstermekte ayrıca durağan olmayan sinyallerin analizinde rahatlıkla kullanılabilir. Bu nedenle DD sinyal işleme çalışmalarında sıklıkla kullanılan bir teknik haline gelmiştir (Vetterli ve Kovacevic, 1995; Merry, 2005; Xizheng ve ark, 2010).

DD, FD'yi temel almasına rağmen FD'de olduğu gibi sabit fonksiyonları kullanmak yerine, EEG gibi ani değişim gösteren sinyaller için zaman-frekans çözünürlüğünü sağlayabilecek, frekansa bağlı büyüklüğü değişebilen ve ayarlanabilen pencere fonksiyonlarını kullanmaktadır. Bu pencereler düşük frekanslar için geniş, yüksek frekanslar için dar olacak şekilde değişen pencere boyutlarına sahiptir. Böylece bütün frekans aralıklarında optimum zaman-frekans çözünürlüğünü sağlayabilmektedir (Misiti ve ark, 1996; Merry, 2005).

EEG sinyallerine dayanan çalışmaların çoğu hasta ve sağlıklı bireylere ait EEG sinyallerinin karakteristik özelliklerini temsil edecek özniteliklerin çıkartılması ve elde edilen öznitelikler kullanılarak hasta ve sağlıklı bireylerin sınıflandırılması olmak üzere iki aşamadan oluşmaktadır. EEG sinyallerinde öznitelik çıkartma ve sınıflandırma için kullanılan birçok istatistiksel yöntem vardır. EEG sinyallerinden hızlı ve yüksek doğrulukta bilgi elde etmek için mevcut ve sürekli geliştirilmekte olan farklı boyut indirgeme ve sınıflandırma yöntemlerinin performanslarının karşılaştırılması istatistikte ilgi duyulan güncel bir konudur.

EEG sinyalleri epilepsi teşhisine yönelik çalışmalarda sıklıkla kullanılmıştır. Literatüre bakıldığında EEG sinyallerini kullanarak yapılan çoğu çalışmada, epilepsi hastası ve sağlıklı bireylerin ayırımında oldukça yüksek sınıflandırma başarılarının elde edildiği görülmektedir. Lima ve ark (2010) yaptıkları çalışmada Bonn veri tabanındaki (Andrzejak ve ark, 2001) hasta olmayan ve epilepsi hastası olan bireylerden kriz anında alınan (sırasıyla A ve E veri setleri) EEG sinyallerini kullanmışlardır. Destek Vektör Makinesi (DVM) ile yaptıkları sınıflandırmada %99,9 doğruluk oranı elde etmişlerdir. Subaşı ve Gürsoy (2010) yaptıkları çalışmada, hasta olmayan ve epilepsi hastası olan bireylerden nöbet sırasında alınan EEG sinyallerini (Bonn veri tabanı A-E veri setleri) Ayırık Dalgacık Dönüşümü (ADD) yöntemiyle alt frekans bantlarına ayırmıştır. Elde ettikleri alt frekans bantlarından çıkarttıkları özniteliklerin boyutlarını Temel Bileşenler Analizi (TBA) ve Bağımsız Bileşen Analizi (BBA) yöntemleriyle indirgenmiş ve Destek Vektör Makinesi (DVM) ile sınıflandırmışlardır. BBA ile boyutu indirgenen özniteliklerle %99,5 ve TBA ile boyutu indirgenen özniteliklerle %98,75 doğruluk oranı elde etmişlerdir. Kaya ve ark (2012) yaptıkları çalışmada, Bonn veri tabanı A-E veri setlerini kullanarak epileptik sinyallerin sınıflandırılması için EEG sinyallerinin DD ile spektral çözümlemesini yaptıktan sonra elde ettikleri öznitelik vektörlerini Karar Ağaçları ve Karar Kuralları yöntemlerini kullanarak sırasıyla %96.6 ve %99.70 oranında doğru sınıflandırma başarıları sağlamışlardır. Yapılan bu çalışmalarda elde edilen bu yüksek sınıflandırma başarısının temel nedeni kullanılan istatistiksel yöntemlerden ziyade karşılaştırılan EEG sinyallerinin epilepsi hastalarından kriz anında alınan sinyaller ile sağlıklı bireylerden alınan sinyaller olmasıdır. Çünkü kriz anında kaydedilen sinyaller ile sağlıklı bireylerden kaydedilen sinyaller arasında oldukça belirgin farklılıklar söz konusudur. Bu nedenle karakteristikleri çok farklı olan bu tip sinyallerin uzmanlar tarafından görsel olarak değerlendirilip epileptik ya da normal olduğuna karar verilmesi basit bir işlemdir. Yukarıda söz edilen çalışmaların aksine bu çalışmada kullanılan EEG sinyalleri uzman hekim tarafından epilepsi tanısı konmuş bireylerden kriz geçirmediikleri anda kaydedilen sinyaller ile

normal bireylerden kaydedilen sinyallerdir. Ayrıca EEG sinyallerinde meydana gelebilecek epileptik belirtiler beynin belli bir bölgesinde meydana gelebileceđi gibi aynı anda birden fazla bölgede de gözlenebilir (Adeli ve ark, 2003; Mahajan ve ark, 2011). Bu nedenle bu çalışmada 16 kanaldan alınan kayıtların tamamı kullanıldı.

Bu çalışmada, uzman hekim tarafından incelenerek seçilen ve farklı epileptik belirtiler içeren EEG sinyalleri ile epileptik olmayan EEG sinyallerinden ADD ile sinyalleri temsil edecek özniteliklerin çıkartılıp, TBA ve BBA yöntemleri ile boyutlarının indirgenmesinin sınıflandırma başarısı üzerine etkilerinin belirlenmesi ve LDA ile DVM yöntemlerinin sınıflandırma performanslarının karşılaştırılması amaçlandı.



2. GENEL BİLGİLER

2.1 Beyin

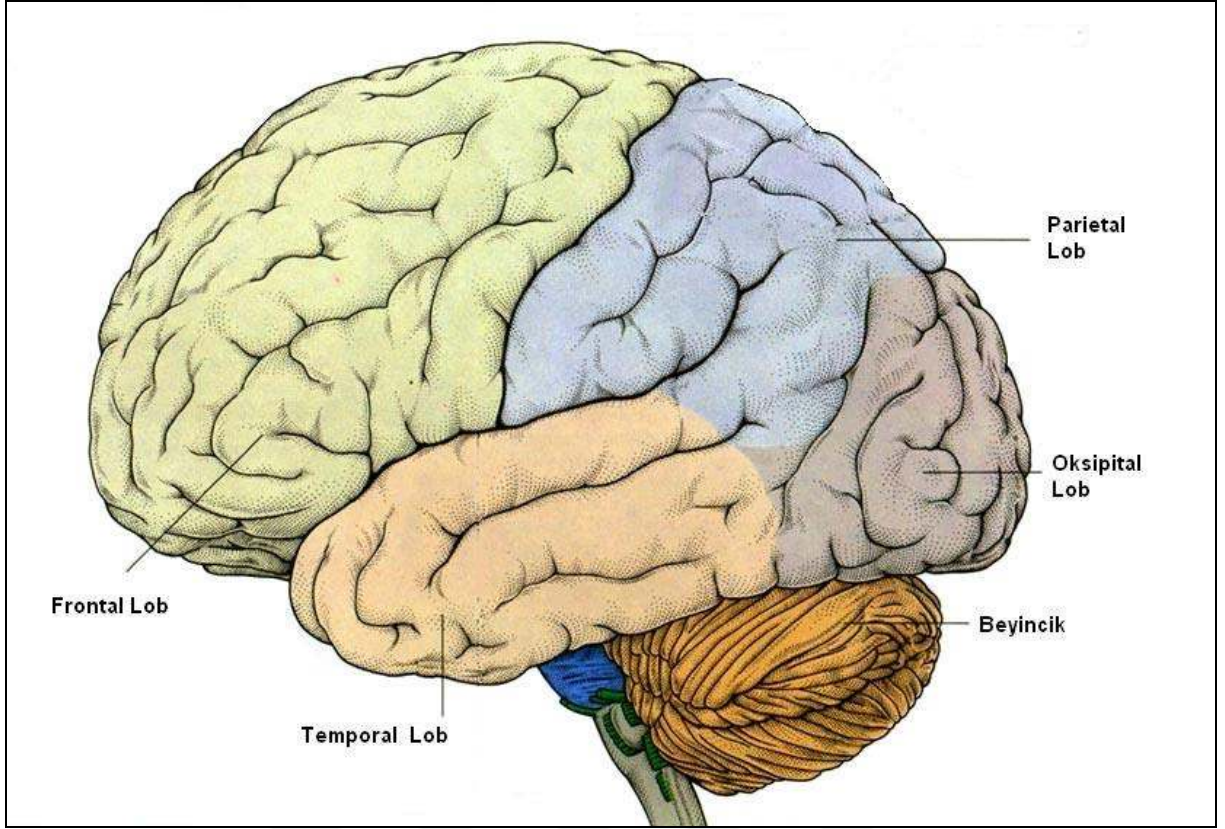
Beyin, nöron adı verilen sinir hücrelerinin gövdeleri ve liflerinin kafatası içinde yoğun bir şekilde toplanması ve çevresini saran nöroglia veya glia adı verilen gri bir maddeden oluşan, karmaşık işlemleri bir arada yapabilen organdır (Yazgan ve Korürek, 1996). İnsan beyinde yaklaşık 100 milyar beyin hücresi (nöron) vardır. Nefes alma, hareket etme, duyu organlarıyla hissetme gibi motor aktiviteleri ve bilişsel faaliyetleri kontrol eden beyin, tüm fonksiyonlarını sinir ve kas sistemine bağlı olarak gerçekleştirir (Ropper, 2009; Ullsperger ve Debener, 2010).

Beynin büyük bir kısmı proteinden oluşur. Beynin çalışması için gerekli olan enerji, glikozun oksijenle yanmasıyla elde edilir. Beyin insan vücudunun sadece %2'sini oluşturmasına rağmen vücudun tükettiği oksijenin yaklaşık %25'ini kullanmaktadır (Yazgan ve Korürek, 1996).

İnsan beyni yapısal olarak frontal (ön) lob, parietal (yan) lob, temporal (şakak) lob, oksipital (arka) lob olmak üzere dört lobdan ve beyincikten oluşmaktadır (Şekil 2.1).

Frontal Lob: Göz kürelerinin üzerinden başlar ve beyin alında saçlı derinin başladığı yere kadar olan kısmına verilen isimdir. Bu bölgenin en belirgin özelliği insanda akıl özelliklerinin olduğu bölüm olmasıdır. Bu lob sayesinde insan dikkatini toplar, kısa süreli belleği kullanabilir, düşünebilir ve koku alma gibi işlevleri yerine getirebilir. Ayrıca problem çözme ve hafıza çalışması esnasında da etkilidir (Ropper, 2009; Ullsperger ve Debener, 2010; Carter, 2014).

Temporal Lob: Görsel ve işitsel bellek, öğrenme ve konuşma gibi işlemlerin yapılmasından sorumlu olan beyin bölgesidir. Hasar görmesi durumunda konuşma ve işitme problemleri meydana gelmektedir. Ayrıca duygusal denge ve sosyalleşme konuları da temporal lobun çalışması sonucu ortaya çıkan özelliklerdendir (Ropper, 2009; Ullsperger ve Debener, 2010; Carter, 2014).



Şekil 2.1 Beyin lobları ve beyincik

Web_1. (2009) <http://www.coheadquarters.com/coOuterBrain1x.jpg>'dan modifiye edilmiştir.

Oksipital Lob: Kafatasının arka kısmında yer alan bu lob, görme işlemlerinden sorumlu olan beyin bölümüdür. Hasar görmesi durumunda insanda görme bozuklukları meydana gelir (Ropper, 2009; Carter, 2014).

Parietal Lob: Beynin üst orta kısmında yer alır. Duygusal bilgileri işler, vücuttan gelen dokunma, ağrı, sıcaklık ve basınç duyularını alarak cevap üretir. Ayrıca okuma ve matematiksel işlemlerde akılda tutmayı sağlar (Ropper ve Brown, 2005).

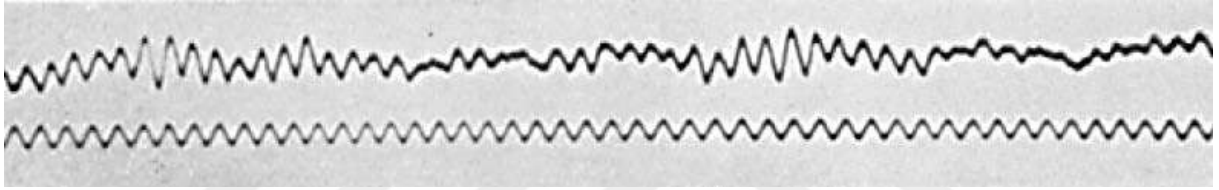
Beyincik: Beynin arka kısmında yer alır. Görevi yürüme işlemi sırasında denge ve düzeni sağlamaktır (Ropper, 2009; Ullsperger ve Debener, 2010; Carter, 2014).

2.2 Elektroansefalogram (EEG)

2.2.1 EEG'nin Tarihçesi

Beyin tarafından üretilen sinyaller üzerinde yapılan ilk önemli çalışma, 1875 yılında İngiliz bilim adamı Richard Caton'ın bir deney hayvanının beynindeki elektriksel aktiviteyi tespit ettiği deneyin bulgularını açıkladığı çalışma olarak kabul edilmektedir (Niedermeyer ve Silva, 2005).

Bundan sonraki ilk önemli gelişme, 1924 yılında, Alman psikolog Hans Berger'in Şekil 2.2'de gösterilen ilk insan EEG'sini kaydetmesidir (Niedermeyer ve Silva, 2005). Bu sinyaller deneklerin kafataslarına bugün kullanılan elektrotlar yerine metal bantlar yerleştirilmesi ve bir galvanometre yardımıyla dalgalanmaların kağıt üzerine aktarılmasıyla elde edilen, zigzag görünümlü sinyallerdir.



Şekil 2.2 Kaydedilen ilk EEG
Niedermeyer ve Silva (2005)'ten modifiye edilmiştir.

Berger, 1924 - 1938 yılları arasında yaptığı deneylerle, insan beyninin düzensiz ve düşük voltajlı elektrik sinyallerini incelemiş ve çalışmalarının ilerlemesiyle beyin dalgalarının tamamen rastlantısal olmadığını, zamanla değişen bazı periyodik ve düzenli dalgalanmaların gerçekleştiğini gözlemlemiştir. Gözlemleri sonucu, uyku ile uyanıklık süreçlerinde farklı beyinsel aktivitelerin gerçekleştiğini belirtmiştir. Ayrıca epilepsi nöbeti esnasındaki beyin sinyallerini gözlemleyerek bu dalgaların insanların patolojik durumları hakkında bilgi verebileceğini söylemiştir.

Adrian ve Matthews'un Berger'in çalışmalarına yaptıkları katkı, ölçülen sinyalleri güçlendirerek 10 -12 Hz arasında dalgaları kaydetmeyi başarmış olmalarıdır. Adrian ve Matthews'un alfa ritmi ismini verdikleri bu dalgaların, deneklerin görme alanındaki bir objeye ilgilerini yöneltmeleri durumunda kaybolduklarını gözlemlemiştirler (Bronzino, 1999).

Kornmüller daha geniş beyin bölgelerinin incelenebilmesi için daha fazla sayıda elektrot kullanılmasının ve çok kanallı kayıt yapılmasının ne derece önemli olduğunu ilk fark eden kişi olmuştur. Epileptik dikensi yapıların gösterimi Fisher ve Löwenbach tarafından gerçekleştirilmiştir (Sanei ve Chambers, 2007).

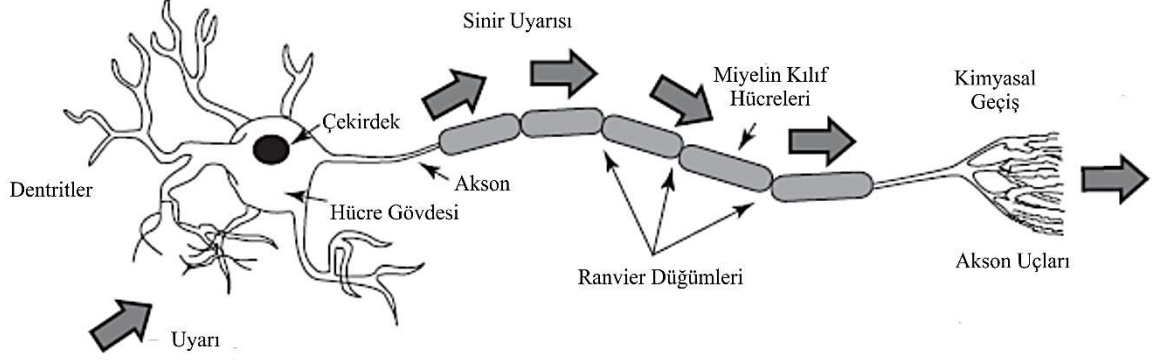
20. yüzyıldan günümüze kadar gelen EEG araştırmalarının sayısı hızlı bir şekilde artmış ve EEG teknolojisi sürekli gelişmiştir. EEG sinyalleri artık yalıtılmış odalarda, daha iyi kalibre edilmiş sistemlerle, çok daha az gürültü barındıran sinyaller olarak ölçülmekte; açık beyin ameliyatlarından mikro boyuttaki nöron araştırmalarına, beyin bilgisayar ara yüzlerine kadar geniş bir alanda kullanılmaktadır (Ullsperger ve Debener, 2010; Carter, 2014).

2.2.2 EEG'nin Tanımı

EEG, beyin faaliyetleri sırasında oluşan elektriksel dalgalanmaların kafatası üzerine yerleştirilen elektrotlar aracılığıyla kaydedilmesidir (Nunez ve Srinivasan, 2006). Beynin nörolojik fonksiyonları ve nöropsikolojik özelliklerinin anlaşılması ve araştırılması, ayrıca teşhis ve tedaviye yönelik birçok klinik amaç için günümüzde EEG sinyallerine sıklıkla başvurulmaktadır (Ullsperger ve Debener, 2010).

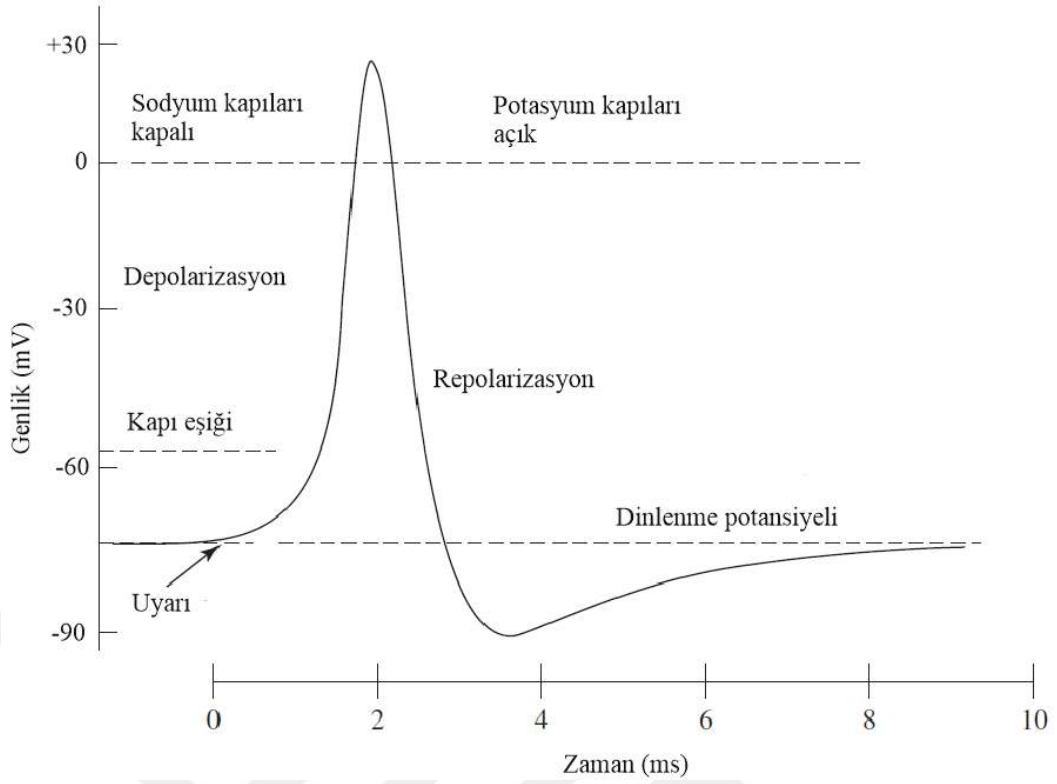
2.2.3 EEG Sinyallerinin Oluşumu ve Ölçümü

Sinir hücrelerinde bilgi iletimi elektriksel potansiyeller ile sağlanır. Bu elektriksel potansiyellere aksiyon potansiyeli adı verilir. Aksiyon potansiyeli, hücre zarındaki iyon geçişleri ile oluşur. İyon geçişlerinin başlaması için dışarıdan bir uyarı gelmesi gereklidir. Eğer uyarı eşik seviyesinin üstünde bir değere ulaşırsa hücre zarında iyon geçişleri meydana gelir. Hücrede oluşan uyarı, dentritlerden başlayarak, hücre gövdesi, akson, Ranvier düğümleri ve akson uçlarına doğru ilerler (Sanei ve Chambers, 2007). Şekil 2.3'te insandaki sinir hücrelerinin genel yapısı gösterilmiştir.



Şekil 2.3 Sinir hücresinin genel yapısı
Sanei ve Chambers (2007)'den modifiye edilmiştir.

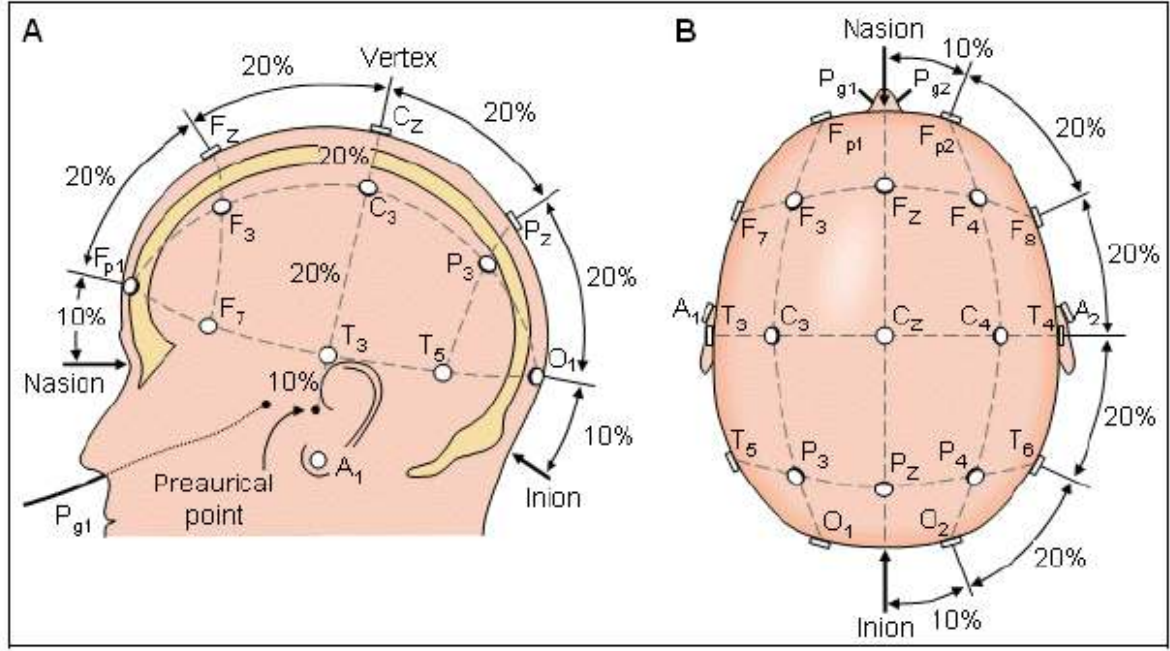
Sinir hücresinin uyarıldığı durumda meydana gelen genlik değişimi Şekil 2.4'te görülmektedir. Şekilde görüldüğü gibi hücre normalde negatif yüklüdür ve yaklaşık -70 mV genliğine sahiptir. Hücrenin dendritleri uyarıldığı zaman, yeteri kadar uyarı oluşturulduğu takdirde (eşik seviyesini aşacak kadar uyarı olduğu durumda), hücre içi iyonlarında bir geçiş söz konusu olur. Öncelikle sodyum kapıları açılır ve hücre negatif durumdan pozitif duruma geçmeye başlar, bu sürece depolarizasyon denir. Daha sonra ise potasyum kapıları açılmaya başlar ve kısa bir süre sonra hem sodyum hem de potasyum kapıları açılmış olur. Böylece bir başka aksiyon potansiyelinin oluşumu engellenir. Daha sonra ise hücre negatif duruma geçmeye başlar. Hücre, dinlenir durumda olduğu halden daha negatif bir hale geçer (yaklaşık -90 mV'a kadar düşer), fakat kısa süre sonra dinlenme durumundaki potansiyeline yeniden kavuşur, bu sürece repolarizasyon denir (Sanei ve Chambers, 2007).



Şekil 2.4 Aksiyon potansiyeli
Sanei ve Chambers (2007)'den modifiye edilmiştir.

EEG sinyali ise, serebral korteksteki sinir hücrelerinde uyarı oluştuğunda meydana gelen potansiyel farktan dolayı dentritlerde oluşan akımı ifade eder. Bu akım kafatasında elektriksel bir alan oluşturarak, EEG cihazıyla ölçülebilir hale gelir (Sanei ve Chambers, 2007).

EEG ölçümü genelde kafatasında saçlı deriye yerleştirilen elektrotlar yardımıyla yapılır. Bu ölçüm sırasında kişi acı hissetmez ve vücuduna elektrik verilmez. Elektriksel potansiyeller, beyindeki sinir hücrelerinde gerçekleşen elektriksel aktivitelerin bir toplamı şeklinde elektrotlar tarafından algılanır. Elektrotlar kafatasında saçlı deriye Şekil 2.5'te gösterildiği gibi 10-20 Sistemi adı verilen uluslararası bir standarda göre yerleştirilir (Malmivuo ve Plonsey, 1995).



Şekil 2.5 Standart 10-20 elektrot yerleşimi
Malmivuo ve Plonsey (1995)'ten modifiye edilmiştir.

EEG sinyallerinin klasik değerlendirilmesinde frekans ve genlik kullanılır. Kafa üzerinden kaydedilen EEG sinyallerin genliği tepeden tepeye 1–100 μ V ve frekans bandı 0,5–100 Hz'dir (Yazgan ve Korürek, 1996). Beynin aktivite düzeyi yükseldikçe EEG sinyallerinin frekansı artar, buna karşılık genlik küçülür. Dolayısıyla EEG sinyallerinde frekans – genlik büyüklükleri ters orantılı olarak değişmektedir. EEG sinyalleri birçok alt frekans bileşenlerini içerse de belli frekans aralıklarında belli bir karakteristik özelliğe sahip özel EEG sinyalleri gözlemlenir. Bu nedenle EEG sinyalleri frekans bant aralıklarına göre sınıflandırılır.

Beynin elektriksel aktivitelerinde meydana gelen bazı değişiklikler ve düzensizlikler birtakım hastalıkların habercisi olabilmektedir. Bu hastalıklardan bir tanesi de halk arasında sara hastalığı olarak bilinen epilepsidir.

2.3 Epilepsi

Epilepsi, beyinde bulunan sinir hücrelerinde nöbet (atak) adı verilen olağan dışı ve geçici elektrik boşalmaları sonucu meydana gelen, vücutta kontrol edilemeyen titreme, morarma, nefes darlığı, bayılma gibi istem dışı hareketlere yol açan nörolojik bir hastalıktır (Ropper ve Brown, 2005; Ropper, 2009).

Epilepsi teşhisine yönelik yapılan çalışmalarda, EEG sinyalleri kapsadıkları frekans aralıklarına göre dört alt frekans bandına ayrılmıştır, bunlar: delta (δ), teta (θ), alfa (α), beta (β) bantlarıdır. Tablo 2.1’de EEG sinyallerinin içerdikleri alt frekans bantları ve bu bantlara verilen özel isimler gösterilmektedir (Adeli ve ark, 2003; Sanei ve Chambers, 2007; Mahajan ve ark, 2011).

Tablo 2.1 EEG sinyallerinin frekans bantları

Bant Adı	Frekans Aralığı
Delta (δ)	0,5 – 4 Hz
Teta (θ)	4 – 7,5 Hz
Alfa (α)	8 – 13 Hz
Beta (β)	13 – 30 Hz

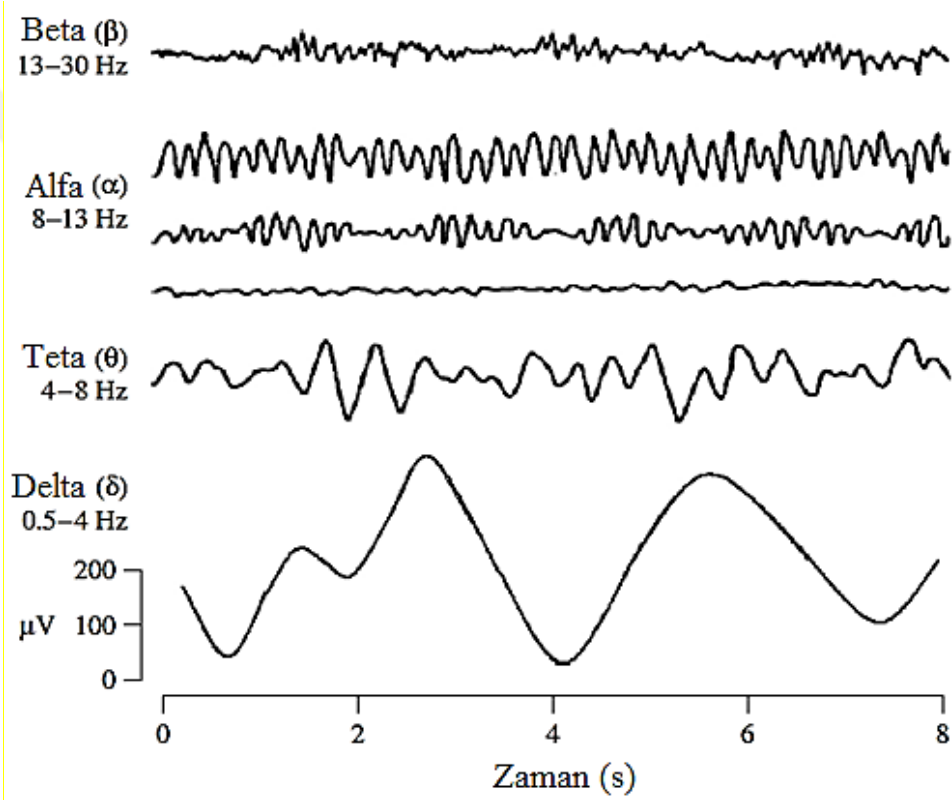
Delta (δ) bandı: Frekansları 0.5-4 Hz ve genlikleri 100 mikrovolt (μV)’tan küçüktür (Sanei ve Chambers, 2007). Süt çocukları, derin uyku ve genel anestezi durum gibi beyin çok düşük aktivite gösterdiği durumlarda gözlemlenir. Özellikle yetişkin insanlarda delta dalgasının sık gözlenmesi ciddi beyinsel problemlerin habercisi olabilir (Ullsperger ve Debener, 2010).

Teta (θ) bandı: Frekansları 4-7,5 Hz arasında değişir (Sanei ve Chambers, 2007). Özellikle çocuklarda çok sık görülür. Yetişkinlerde ise rüyalı uyku, orta derinlikte anestezi durum gibi beyin düşük aktivite gösterdiği durumlarda ortaya çıkar. Bu dalganın çocuklarda görülmesi normal iken yetişkinler için anormaldir (Klimesch, 1999; Mitchell ve ark, 2008). Özellikle yetişkin insanlarda teta dalgasının sık gözlenmesi (tıpkı delta dalgalarında olduğu gibi) ciddi beyinsel problemlerin habercisi olabilir (Sanei ve Chambers, 2007).

Alfa (α) bandı: Frekansları 8-13 Hz ve genlikleri 50 μV ’dan küçüktür (Sanei ve Chambers, 2007). Dalga şekli sinüzoidal şekle benzemektedir. Alfa dalgaları daha çok yetişkinlerde olmak üzere bütün yaş guruplarında görülmektedir. Alfa dalgaları yetişkinlerde uyanık, gerilimsiz ve beyin aktivitesinin yoğun olmadığı durumlarda kendiliğinden oluşur. Bu dalgalar, yoğun bir şekilde oksipital bölgede ve gözler kapalı olduğu durumlarda daha net gözlemlenir. Dikkatin herhangi bir şeye yöneltilmesi durumunda, alfa dalgaları yerini daha

yüksek frekanslı fakat düşük genlikli beta dalgalarına bırakır (Klimesch, 1999; Ullsperger ve Debener, 2010).

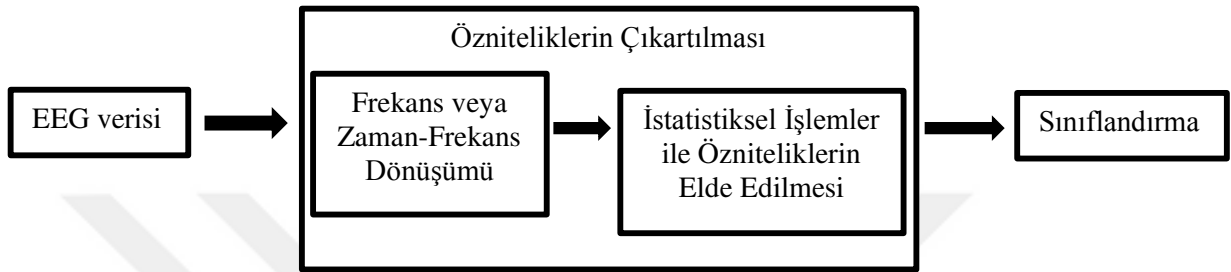
Beta (β) bandı: Frekansları 13-30 Hz ve genlikleri 30 μV 'dan küçüktür (Sanei ve Chambers, 2007). Beynin aktivite düzeyinin yükseldiği sıralarda gözlemlenir. Bireylerin çok stresli olduğu, dikkatlerinin dağıldığı ya da yoğun bir zihinsel iş yürüttükleri zamanlarda görülür. Bu durum saçlı derinin parietal ve frontal bölgelerinde belirgin olarak kaydedilebilir (Sanei ve Chambers, 2007). Şekil 2.6'da EEG sinyallerinin beta, alfa, teta ve delta bantları gösterilmektedir.



Şekil 2.6 EEG sinyallerinin beta, alfa, teta ve delta bantları
Sanei ve Chambers (2007)'den modifiye edilmiştir.

2.4 Epilepsinin Bilgisayar Destekli Teşhisi

Bilgisayar destekli otomatik karar sistemlerinin gittikçe daha güvenilir sonuçlar üretmesi sayesinde EEG sinyallerinin de otomatik olarak yorumlanması üzerine birçok uygulama geliştirilmiştir (Übeyli 2008, 2009; Ocak 2009; Altunay ve ark, 2010). Bilgisayar destekli karar sistemlerinde veri işleme süreci Şekil 2.7’deki blok diyagramıyla özetlenebilir.



Şekil 2.7 Bilgisayar destekli karar sistemlerinde süreç

Şekil 2.7’de gösterilen bu süreç sağlıklı bireyler ile epilepsi hastalarından alınan EEG sinyallerinin bir bilgisayar sistemine girilmesi, bu sinyallerin doğru belirlenmiş özniteliklere dönüştürülmesi ve son olarak bu özniteliklerin sınıflandırılması adımlarından oluşur. EEG sinyallerinin sınıflandırılmasında birçok başarılı teknik geliştirilmiş olmasına rağmen bu tekniklerin başarısı bu yöntemlere giriş olarak verilen öznitelik vektörlerine de bağlıdır. Dolayısıyla sınıflandırmada kullanılan yöntem kadar özelliklerin elde edileceği yöntemin de doğru seçilmesi gereklidir. EEG sinyallerinden özniteliklerin elde edilmesi iki aşamalı olarak ele alınabilir. Birinci aşamada EEG sinyalleri zaman, frekans veya zaman-frekans boyutunda analiz edilir ve sinyalin içerisindeki ilgili örüntüler açıkça görülecek şekilde ayrıştırılmaya çalışılır. Diğer aşamada ise sinyallerin ayrıştırılmış katsayıları üzerinde tercih edilen tekniğe dayalı olarak anlamlı özellikler çıkartılır. Zaman boyutunda ham sinyalin, frekans boyutunda genellikle güç spektrumunun ve zaman-frekans boyutunda ise ayrıştırılmış katsayıların üzerinde ortalama, standart sapma, minimum, maksimum, entropi ve ortalama güç gibi istatistikler hesaplanarak öznitelikler elde edilir. Bu sayede bir EEG sinyali, zamana bağlı bir vektörden zamandan bağımsız birkaç skalere dönüştürülmüş olur. Bazen elde edilen öznitelikler veriyi o kadar iyi temsil eder ki kompleks bir sınıflandırma yöntemi yerine basit bir eşik değeri belirlemek bile sınıfları ayırt edebilmek için yeterli olur. Bu şekilde basit bir sınıflandırma yöntemi kullanarak yakalanan yüksek başarı, genellikle güçlü bir öznitelik

çıkartma aşamasının ürünü olarak yorumlanır (Bronzino, 1999; Merry, 2005; Sanei ve Chambers, 2007; Orhan, 2011).

2.5 Sinyal İşleme Yöntemleri

EEG sinyalleri beynin fonksiyonları ile ilgili çok miktarda bilgi içermektedir. Ancak ham EEG sinyallerinin tıbbi araştırmalarda ve hastalık tanısında kullanılabilmesi için çeşitli sinyal işleme yöntemleri kullanılmalıdır. Sinyallerin analizi, sinyalin kendisini veya sinyali temsil eden özellikleri zaman, frekans ya da zaman-frekans gibi boyutlara taşınarak yapılır. Burada amaç; işlenmemiş verileri bilgi kaybı en az olacak şekilde bu boyutlardan birisine taşıyarak detaylı bilgilere ulaşmaktır (Sanei ve Chambers, 2007).

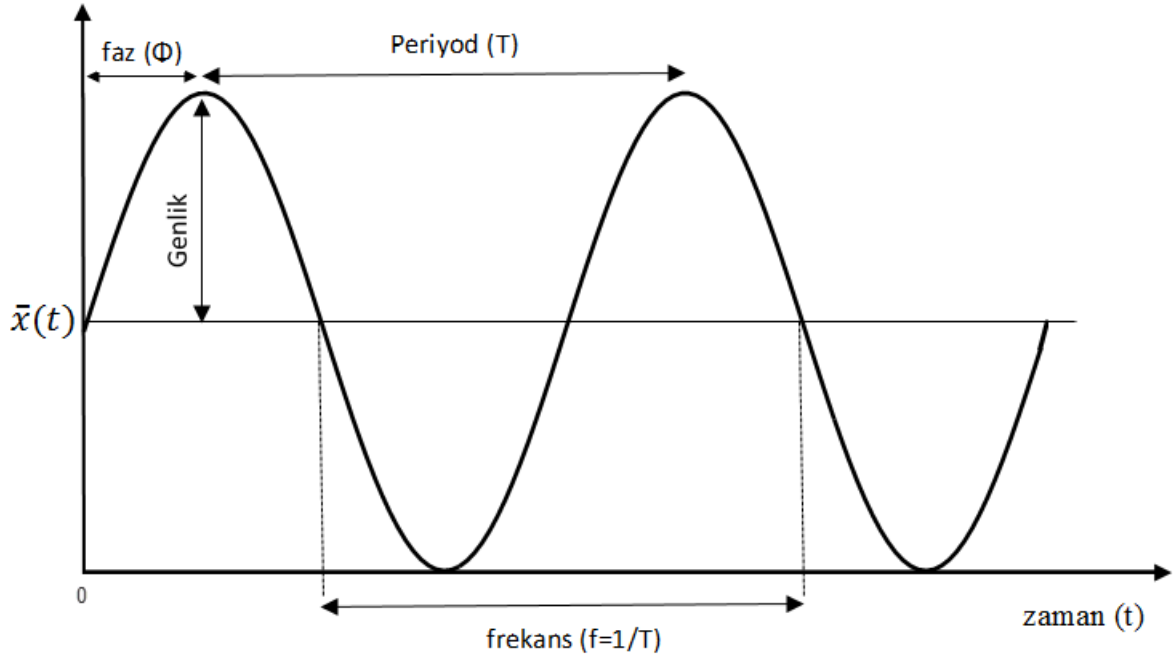
EEG sinyallerinin zamana karşı gerilim grafiği olduğu düşünülürse herhangi bir işlem yapmadan ham EEG sinyalleri üzerinde yapılacak her analizin en temel yöntem olacağı söylenebilir. Sinyallerin yorumlanmasında en basit yöntem, analiz edilecek sinyallerin karakteristiklerine uygun istatistiklerin hesaplanması ve bu istatistiklerin incelenmesidir. Önerilmiş birçok istatistik olsa da EEG sinyalleri için en sık kullanılan istatistikler ortalama, standart sapma, minimum, maksimum, entropi ve ortalama güç değerleridir (Kandaswamy ve ark, 2004; Li ve Wen, 2009; Subaşı ve Gürsoy, 2010; Ersöz ve Özşen, 2011; Ahammad ve ark, 2014). EEG sinyallerindeki epileptik aktivitelerin değişim hızları ve belirginliğine bağlı olarak zaman boyutunda yapılan analizlerde de çok yüksek sonuçlar alınabilir. Örneğin epileptik nöbet anında alınan kayıtlarla nöbet dışında alınan kayıtların yalnızca zaman boyutunda analizinin yapılması iki durum arasındaki farkı ortaya koymada yeterli olabilir. Fakat tüm epilepsi hastalarının sağlıklı kişilerden ayırt edilmesi gibi zor sınıflandırmalar için zaman boyutunda analiz yeterli olmaz ve bu yüzden spektral analiz yöntemleri olan frekans veya zaman-frekans boyut analizleri tercih edilir.

2.5.1 Spektral Analiz

Sonlu uzunluktaki durağan bir sinyalin toplam gücünün frekans üzerindeki dağılımının incelenmesine *Spektral Analiz* denir (Kay ve Marple, 1981; Kay, 1988).

2.5.1.1 Periyodisite

Belirli bir zaman aralığında olayların benzer şekilde gözlenmelerine *periyodisite* denir. T zaman periyodunda kendini tekrarlayan bir serinin gözlem değerleri ortalama etrafında zamana göre düzenli azalış ve artışlar gösterir (Türe, 1988). Şekil 2.8’de periyodik bir fonksiyonun grafiği ve parametreleri verilmiştir.



Şekil 2.8 Periyodik bir fonksiyonun grafiği ve parametreleri

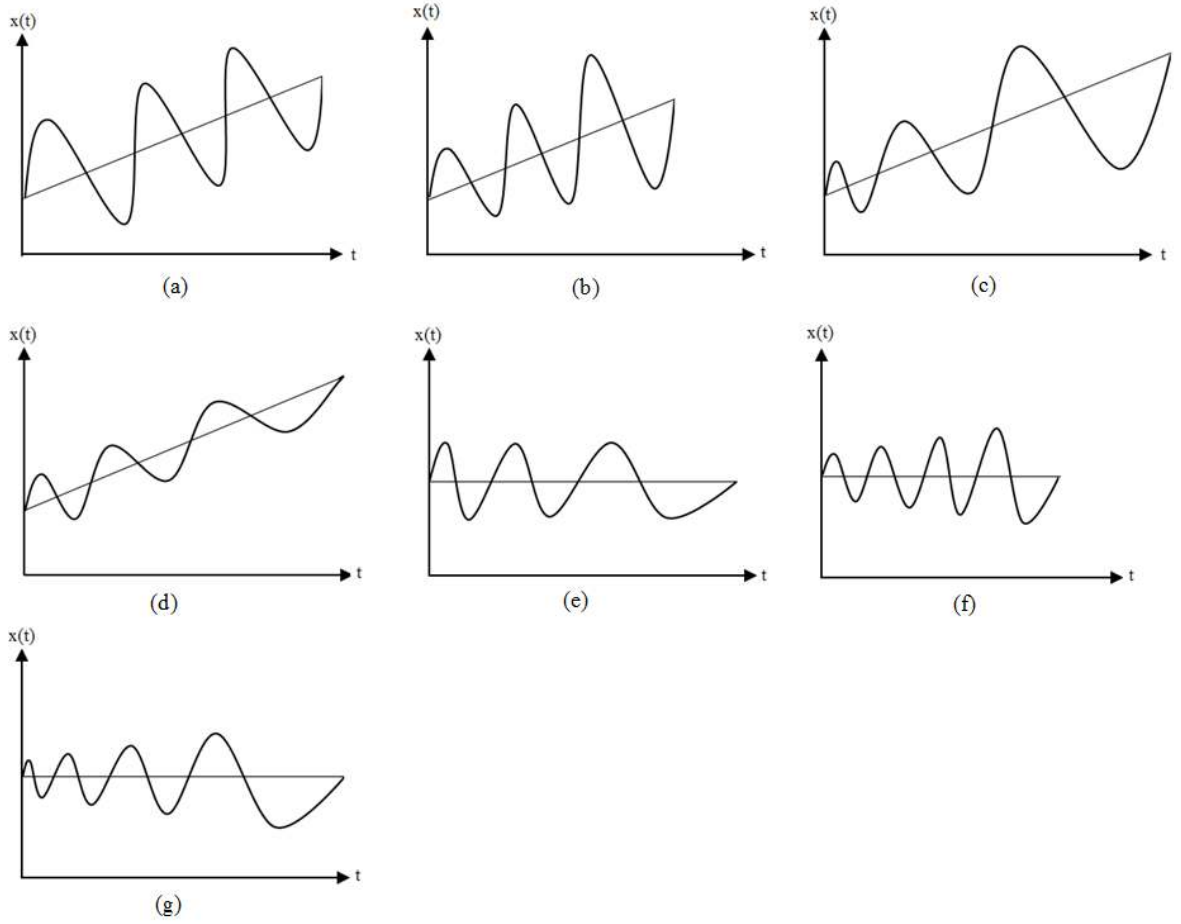
Periyod (Dalga Boyu): Dalganın ortalamaya eşit olduğu noktadan başlayarak bir tepe bir de dip noktası yaptıktan sonra ortalamaya eşit oluncaya kadar geçen zamana (uzunluğa) periyod denir. Diğer bir anlatımla ardışık iki tepe ya da iki dip noktası arasındaki uzunluğa periyod denir (Türe, 1988). Eğer Şekil 2.8’deki fonksiyon bir zaman serisi (sinyal) ise periyod, dalgalanmanın tamamlanması için geçen süre olacaktır.

Frekans: Birim zamana düşen dalga sayısı, fonksiyonun (sinyalin) frekansı olarak tanımlanır (Türe, 1988). Periyod T ve frekans f ile gösterirse $f = 1/T$ olur.

Genlik: Bir dalganın tepe ya da dip noktalarının ortalamaya olan uzaklığına genlik (amplitüd) denir (Türe, 1988).

Faz: Dalganın ilk tepe ya da dip noktasının dalganın başlangıç noktasına olan uzaklığına faz denir (Türe, 1988).

Durağanlık: Zamana göre genliği, periyodu ve ortalaması değişmeyen seriler *Durağan Seriler* olarak adlandırılır. Zamana göre genliği, periyodu ya da ortalaması değişen seriler ise *Durağan Olmayan Seriler* olarak adlandırılır. Şekil 2.8’de verilen periyodik fonksiyon durağan olup zamana göre karakteristik özellikleri değişmemektedir. Şekil 2.9’da durağan olmayan bir zaman serisinin alabileceği 7 durum görülmektedir (Türe, 1988).



Şekil 2.9 Durağan olmayan bir zaman serisinin alabileceği durumlar

Durağan sinyallerin frekans detaylarını analiz etmede oldukça etkin bir yöntem olan Fourier Dönüşümü kullanılır (Merry, 2005).

2.5.1.2 Fourier Serisi ve Fourier Dönüşümü

19. yy. da Fransız bilim adamı Joseph Fourier periyodik bir sinyalin sonsuz sayıdaki sinüzoidaller veya karmaşık üsteller gibi harmonik fonksiyonların doğrusal ağırlıklı toplamı ile ifade edilebileceğini belirtmiştir (Oppenheim ve ark, 2014). Periyodu 2π olan bir fonksiyonun Fourier Serisi eşitlik (2.1) ya da eksponansiyel formda eşitlik (2.2)'deki gibi ifade edilir (Türe, 1988).

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nt) + b_n \sin(nt)] \quad (2.1)$$

$$x(t) = \sum_{k=-\infty}^{\infty} c_k e^{ikt} \quad (2.2)$$

Eşitlik (2.1)'de a_0 , a_n ve b_n Fourier katsayılarını temsil etmektedir ve bu katsayılar aşağıda verilen eşitlikler kullanılarak hesaplanır.

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} x(t) dt$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \cos(nt) dt$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \sin(nt) dt$$

Zaman uzayında tanımlı durağan bir sinyalin frekans uzayına taşınması ve sinyale ilişkin frekans bileşenlerinin ortaya çıkarılması için kullanılan en eski ve klasik yöntem FD'dir (Merry, 2005).

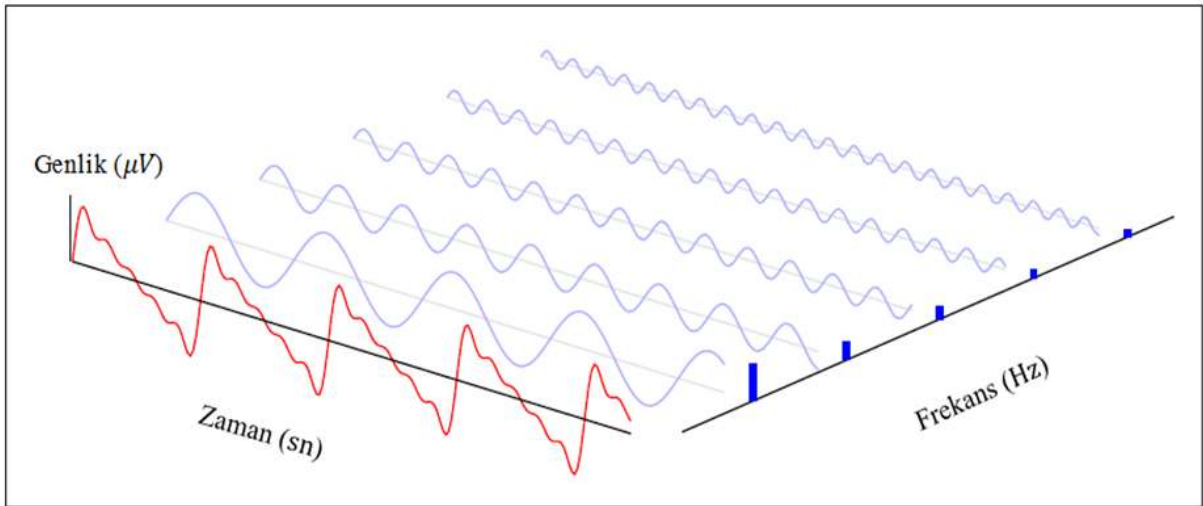
Periyodu T olan bir sinyalin FD'si ve ters dönüşümü sırasıyla eşitlik (2.3) ve (2.4)'teki gibi ifade edilir.

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-i2\pi ft} dt \quad (2.3)$$

$$x(t) = \int_{-\infty}^{+\infty} X(f)e^{i2\pi ft} df \quad (2.4)$$

Eşitliklerde $x(t)$ FD'si alınacak sinyal, $X(f)$ ise FD'si alınmış sinyaldir. Burada t zamanı göstermektedir (Merry, 2005; Oppenheim ve Verghese, 2015).

Kısaca FD ile bir sinyal içerdiği frekans bileşenleri ve ağırlıkları cinsinden tanımlanır. Böylelikle zaman uzayında tanımlı bir sinyalin frekans uzayına taşınması ile doğrudan gözlemlenmesi mümkün olmayan frekans bileşenlerinin görülebilmesi sağlanır. Şekil 2.10'da görüldüğü gibi FD basit bir şekilde zaman bölgesinde bulunan periyodik bir sinyalin frekans bölgesine taşınması olarak tanımlanabilir (Merry, 2005).



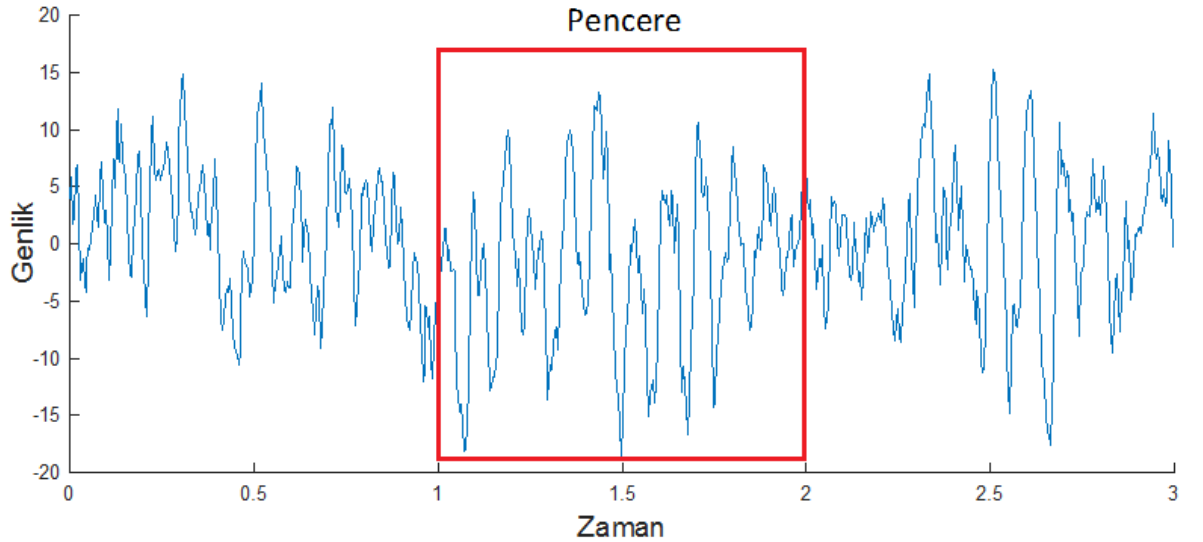
Şekil 2.10 Fourier Dönüşümü

Sinyallerin frekans detaylarını analiz etmede oldukça etkin bir yöntem olan Fourier Dönüşümü ile sinyal zaman bölgesinden frekans bölgesine taşındığı için zaman bilgisi kaybolur. FD, periyodik sinyallerde sabit frekans bileşenlerini başarıyla tespit ederken durağan olmayan sinyallerde zaman bilgisini yok etmesi nedeniyle sinyaldeki tüm değişimleri

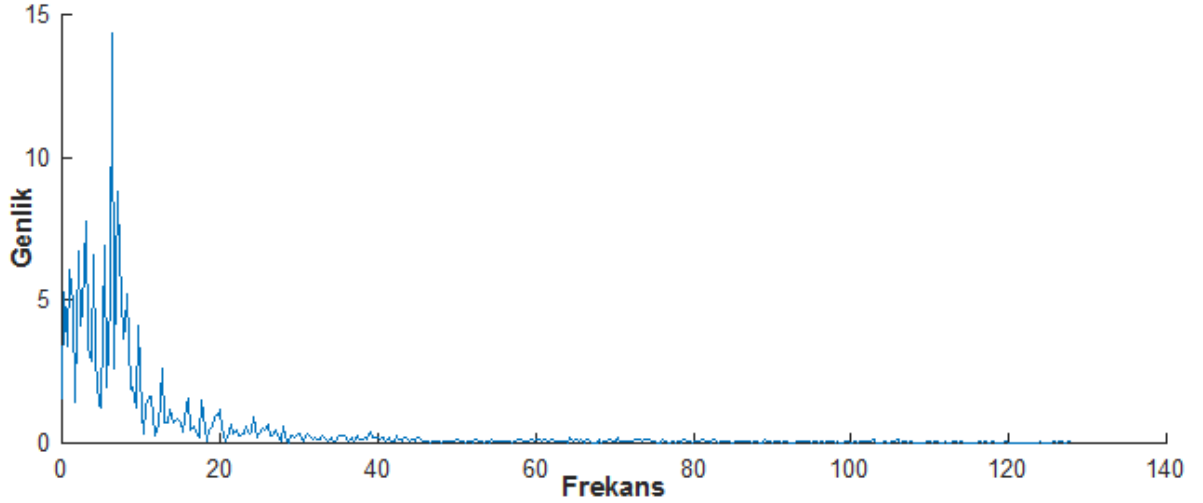
iyi belirleyememektedir (Merry, 2005). Pratik uygulamalarda en sık karşılaşılan sinyallerin durağan olmayan yapıya sahip olması nedeniyle sinyal analizine sadece frekans boyutundan yaklaşmanın doğru olmadığı söylenebilir. EEG sinyalleri gibi durağan olmayan rastgele sinyallerde değişimlerin zamanın hangi anında olduğunu yakalamak ve bununla beraber bir frekans analizi yapabilmek için Kısa Zamanlı Fourier Dönüşümü (KZFD) ya da Dalgacık Dönüşümü (DD) kullanılabilir.

2.5.1.3 Kısa Zamanlı Fourier Dönüşümü

Denis Gabor, 1946 yılında Pencereleme Yöntemi'ni kullanarak, sinyalin küçük bir parçasını zaman-tanım aralığında ele almış, sinyali zaman ve frekansın fonksiyonu olarak iki boyutta ifade etmiştir. KZFD ile orijinal sinyal sabit uzunluklu pencerelere bölünür ve bu parçaların her birine ayrı ayrı FD uygulanır. Bu pencere içerisinde sinyal durağan kabul edilir ve pencere sinyalin sonuna kadar kaydırılarak dönüşüm tamamlanır. Şekil 2.11'de zaman bölgesindeki bir sinyalin pencerelenen 1s'lik kesiti ve Şekil 2.12'de bu pencerelenen kesite uygulanan FD gösterilmektedir.



Şekil 2.11 Sinyalin pencerelenmesi



Şekil 2.12 Pencerelenen sinyale uygulanan Fourier Dönüşümü

KZFD ile mükemmel frekans çözünürlüğü yerine daha düşük frekans çözünürlüğü ve makul derecede zaman çözünürlüğü elde edilmektedir (Merry, 2005).

KZFD’de alçak ve yüksek frekans bileşenleri için pencere genişliğinin değiştirilmesi mümkün değildir. Sadece farklı pencere genişlikleri ile KZFD birkaç kez tekrarlanarak sinyalin birkaç zaman-frekans gösterimi elde edilebilir. Ancak pratik uygulamalarda karşılaşılan birçok sinyal için frekans boyutunda kullanılan pencere uzunluğunun esnek seçilebilmesi önemlidir. Hem zaman hem de frekans boyutunda pencere uzunluğunun ayarlanabildiği, EEG sinyalleri gibi durağan olmayan sinyallerin zaman-frekans boyutunda analiz edilmesini sağlayan DD, uygulamalarda gösterdiği başarılı sonuçlar neticesinde son yıllarda EEG sinyallerinin analizinde popüler bir yöntem haline gelmiştir (Adeli ve ark, 2003; Ocak, 2009; Mahajan ve ark, 2011). Bu çalışmada EEG sinyallerinin analizi DD ile gerçekleştirilmiştir.

2.5.1.4 Dalgacık Dönüşümü

Dalgacık terimi ilk olarak 1909 yılında Alfred Haar’ın çalışmasında geçmiştir. Haar’dan sonra birçok araştırmacı dalgacık teorisi üzerine yoğun bir şekilde çalışmıştır. Dalgacık teorisinin gerçek uygulama alanlarında kullanılarak yaygınlaşması ise ancak seksenli yıllarda olmuştur. Dalgacık teorisi ilk olarak sismik sinyallerin analizinde uygulanmıştır (Morlet ve ark, 1982). Daha sonra bunu Ingrid Daubechies’in (1988) yaptığı teorik çalışma takip etmiştir. Dalgacık Dönüşümü (DD) için kullanılan fonksiyonlar içerisinde en düzenli olanı ve bu

yüzden de en çok kullanılanı Daubechies dalgacığı olmuştur. Aslında Haar Dalgacığı'nın aynısı gibi görünse de Daubechies'in birçok çeşidi türetildiği için Haar Dalgacığı'nın genişletilmiş şekli olduğu söylenebilir. Sinyal analizine dayalı yapılmış birçok çalışmada gösterdiği yüksek başarı dolayısıyla Daubechies, uygulamalarda en çok tercih edilen dalgacık fonksiyonudur.

Biyolojik sinyallerin büyük bir çoğunluğu durağan olmayan yani genlik ve frekans bilgisi zamana bağlı değişen sinyallerden oluşmaktadır. KZFD yöntemi ile sinyal pencerelenerek küçük parçalara ayrılmış ve her parçadaki frekans içeriklerinin bulunması ile zamana bağlı değişen spektral analizin yapılabilmesi mümkün olmuştur. KZFD'de kullanılan pencere boyutunun sabit olması nedeniyle, KZFD sinyalin zaman-frekans içeriğini sabit frekans ve zaman çözünürlüğü ile verir. DD'de ise düşük frekanslar için geniş, yüksek frekanslar için dar olacak şekilde değişen pencere boyutlarının kullanılmasıyla bütün frekans aralıklarında optimum zaman-frekans çözünürlüğü elde edilmektedir (Misiti ve ark, 1996; Merry, 2005).

DD sinyali farklı frekanslar için farklı çözünürlüklerde analiz eden bir yaklaşım olduğundan sinyal işlemede kullanılan “zaman-frekans” kavramı, DD'de “farklı zaman-ölçek” kavramına dönüşür (Merry, 2005; Orhan, 2011).

2.5.1.4.1 Sürekli Dalgacık Dönüşümü

Sürekli Dalgacık Dönüşümü (SDD) bir ana dalgacık fonksiyonunun ölçeklenmesi ve ötelenmesi ile oluşturulan fonksiyonlara sinyalin izdüşümünün alınması ile elde edilir. Ana dalgacık fonksiyonu, dalgacıklar üretmek için kullanılan bir prototip fonksiyondur. Ana dalgacık fonksiyonunun ötelenmesi ve ölçeklenmesiyle dalgacık katsayıları türetilir (Vetterli ve Kovacevic, 1995).

Bir $x(t)$ sinyalinin SDD eşitlik (2.5)'teki gibi tanımlanır:

$$W(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t - \tau}{s}\right) dt \quad (2.5)$$

Eşitlik (2.5)'te τ öteleme parametresini, s ölçek parametresini, $\psi(t)$ ana dalgacık fonksiyonunu ve $W(s, \tau)$ dalgacık katsayılarını göstermektedir. Dönüşüm için kullanılan farklı pencere fonksiyonları ana dalgacığın ötelenmesi ve ölçeklendirilmesiyle türetilmektedir. τ öteleme parametresi, pencere fonksiyonunun sinyal üzerindeki konumunu değiştirir.

Böylelikle zaman ekseninde sinyal süresince pencereyi hareket ettirir. Ölçek parametresi s , pencere fonksiyonunu genişletir ya da daraltır. s 'nin büyük değerleri, detay içermeyen genel görünüm, küçük değerleri ise detaylı görünüm için uygundur. $\frac{1}{\sqrt{s}}$ çarpanı ise, normalizasyon çarpanıdır. Bu çarpan sayesinde ana dalgacık fonksiyonu ile bu fonksiyondan elde edilen pencere fonksiyonlarının enerjisinin eşit olması sağlanır (Merry, 2005).

Orijinal sinyal dalgacık katsayılarından eşitlik (2.6)'da verilen ifade kullanılarak tekrar elde edilebilir.

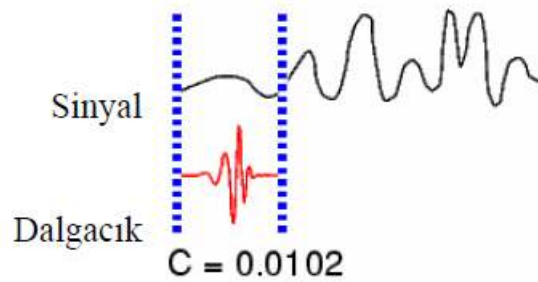
$$x(t) = \frac{1}{C} \int_{-\infty}^{\infty} \int_0^{\infty} \frac{1}{s^2} W(s, \tau) \psi_{s, \tau}(t) ds d\tau \quad (2.6)$$

$$C = \int_0^{\infty} \frac{|\psi(f)|}{|f|} df \quad 0 < C < \infty \quad (2.7)$$

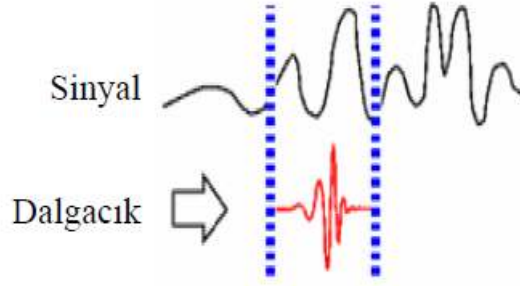
Burada C katsayısı, kabul edilebilirlik sabiti olarak adlandırılır ve eşitlik (2.7)'deki gibi elde edilir. C katsayısı ne kadar büyük olursa sinyal ile ana dalgacık fonksiyonu o oranda benzerdir, denir. $\psi(f)$, dalgacık fonksiyonunun FD'sini ifade etmektedir (Merry, 2005).

Sürekli dalgacık dönüşümü uygulamasında:

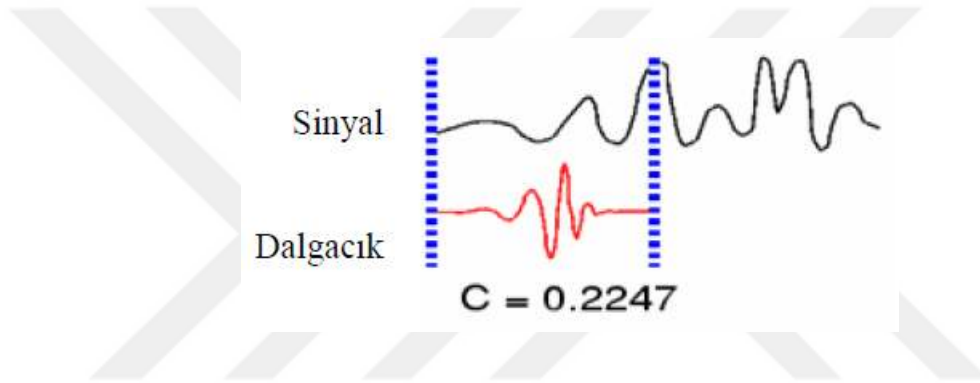
1. Sinyalin karakteristik özelliklerine göre ana dalgacık fonksiyonu seçilir ve orijinal sinyalde seçilen kısımla karşılaştırılır.
2. Sinyalin seçilen bölümüyle dalgacık arasındaki ilişkiyi gösteren C sayısı hesaplanır. C sayısı ne kadar büyük olursa orijinal sinyal ile dalgacık fonksiyonu birbirine o kadar benzer demektir.



3. Dalgacık sağı kaydırılarak sinyalin tamamını kaplayıncaya kadar 1. ve 2. adımlar tekrar edilir.

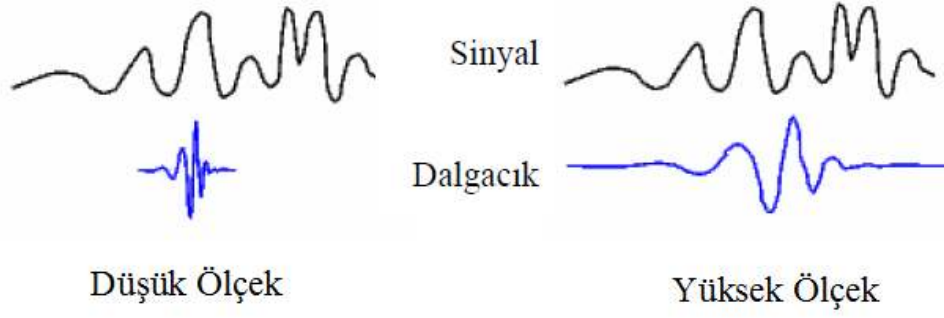


4. Ölçek genişletilerek önceki adımlar tekrar edilir.



5. Bütün ölçek değerleri için adımlar tekrar edilir (Misiti ve ark, 2015).

Bu adımlar sonunda, sinyalin farklı bölümleri ile farklı ölçeklerde üretilen katsayılar elde edilir. Ayrıca dalgacık ne kadar genişletilirse orijinal sinyalin o kadar büyük bir kısmıyla karşılaştırılmış olur (Şekil 2.13). Böylece sinyalin daha kaba özellikleri dalgacık katsayılarıyla ölçülebilir (Misiti ve ark, 1996).



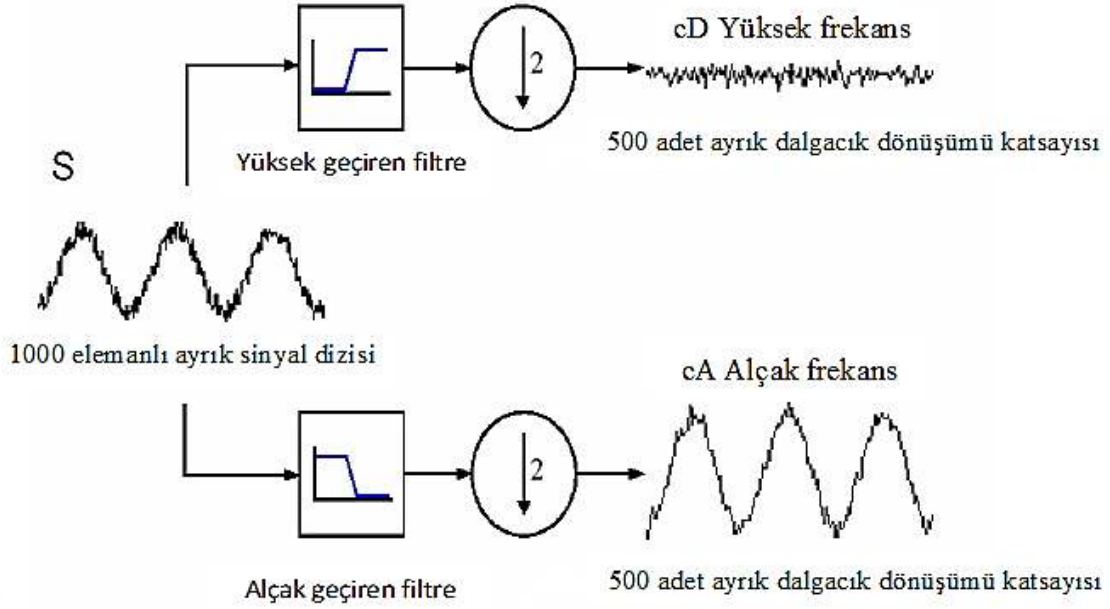
Şekil 2.13 Düşük ve yüksek ölçek
Misiti ve ark (1996)'dan modifiye edilmiştir.

SDD kullanılarak orijinal sinyalin yeniden elde edilmesi işlemi nadiren istenen bir durumdur. Sinyalin yeniden elde edilmesini gerektiren durumlarda genellikle Ayrık Dalgacık Dönüşümü (ADD) kullanılır. Bununla beraber bilgisayarda bir sinyale uygulanan SDD hesaplamasında sürekli sinyaller ayrık hale getirilerek işlenir. SDD yönteminde ölçek ve öteleme parametrelerinin sürekli biçimde değişmesinden ötürü her bir birim için dalgacık katsayılarının hesaplanması zordur ve zaman alır. Bu nedenle pratik uygulamalarda SDD yerine ADD kullanılır (Semmlow, 2009).

2.5.1.4.2 Ayrık Dalgacık Dönüşümü

Dalgacık katsayılarının mümkün olan her ölçek için hesaplanması, sinyal içerisinde gereksiz olan bilgilerin de alınmasına neden olur. Bu da dönüşüm işleminin uzamasına sebep olur (Rioul ve Vetterli, 1991). Eğer seçilecek olan ölçek ve pozisyonları, 2^n 'nin üsleri biçiminde seçilirse yapılan analiz daha etkili ve hızlı olur. Böylelikle istenmeyen verilerin analiz işleminden ayıklanması sağlanır. Bu yöntem ADD olarak adlandırılır (Misiti ve ark, 2004).

ADD'de orijinal sinyale bir alçak geçiren bir de yüksek geçiren olmak üzere iki filtre uygulanır. Dalgacık dönüşümünün SDD'ye göre çok daha hızlı yapılmasını sağlayan bu çok pratik filtreleme algoritması, 1988 yılında Mallat tarafından geliştirilmiştir (Gaps, 1995). Şekil 2.14'te bir S sinyaline filtre uygulanması ve uygulanan filtre sonucu elde edilen alçak ve yüksek frekans bileşenleri verilmiştir.

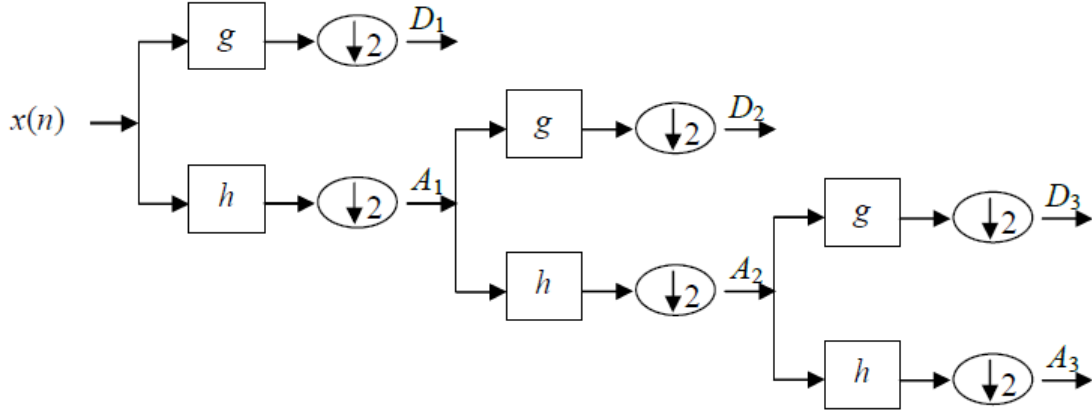


Şekil 2.14 Sinyalin ayrıştırılması
Misiti ve ark (2004)'ten modifiye edilmiştir.

ADD, sinyali alçak ve yüksek geçiren olmak üzere iki ayrı filtreden geçirerek frekans bileşenlerine ayırır. Buradaki alçak frekans bileşenlerine *yaklaşımlar* adı verilir. Yaklaşımlar büyük ölçeklidir ve A ile gösterilir. Yüksek frekans bileşenlerine ise *detaylar* adı verilir. Detaylar küçük ölçeklidir ve D ile gösterilir. Yaklaşım sinyalleri, orijinal sinyali temsil eder ve sinyalin tanımını verir. Detaylar ise sinyalin karakteristiğini ya da ayrıntısını içerir (Misiti ve ark, 2004). Örnek olarak insan sesini ele alırsak; insan sesinden yüksek frekanslar kaldırılırsa konuşmanın içeriği anlaşılabilir. Ancak düşük frekanslar kaldırılırsa içeriği anlaşılmayan, anlam verilemeyen sesler duyulur. DD'de yaklaşımlar kabaca orijinal sinyali temsil eder ve yüksek ölçek ile elde edilir. Detaylar ise düşük ölçekteki bilgilerdir ve sinyalin yüksek frekanslı bileşenleridir. Sinyalin boyutunun değişmemesi için filtre çıkışlarında örnek sayıları eşit olacak şekilde yarıya indirilir. Bu işleme *aşağı örnekleme* denir. Bu işlem, kullanıcı tarafından belirlenen alt bant derinliğine kadar devam eder (Misiti ve ark, 2004).

Hem ADD hem de SDD yönteminde sinyal belli sayıda bileşene ayrıştırılır ve bu işleme *çoklu çözünürlük ayrışımı* denir.

Şekil 2.15'teki diyagramda (g) yüksek ve (h) alçak geçiren filtreler olmak üzere bir $x(n)$ sinyalinin üçüncü seviye ADD yöntemiyle alt bantlarına ayrıştırılması gösterilmiştir.



Şekil 2.15 ADD ile bir sinyalin alt bantlarına ayrıştırılması
Misiti ve ark (2015)'ten modifiye edilmiştir.

Şekil 2.12’de ilk filtreler sayesinde elde edilen sinyaller birinci seviye yaklaşım (A1) ve detay (D1) katsayıları olarak adlandırılır. Sonraki seviyelerin yaklaşım ve detay katsayıları, önceki seviyenin yaklaşım katsayısı kullanılarak hesaplanır. Filtrelerin ardından uygulanan aşağı örnekleme ile sinyalin ayrıştırma sonrası toplam uzunluğu da korunmuş olur (Misiti ve ark, 2015)

ADD için t zaman ifadesi yerine n örnek sırası ifadesi kullanılırken genellikle $s = 2^m$ ve $\tau = k2^m$ olarak tanımlanır. ADD yönteminde bahsedilmesi gereken bir diğer kavram ise ölçek fonksiyonudur. ADD yöntemini etkin bir şekilde kullanabilmek için ilk olarak en uygun çözünürlüğün hesaplanması gerekir. En uygun çözünürlüğün hesaplanmasında orijinal dalgacık biçimi yerine bunun yumuşatılmış bir şekli kullanılır. Orijinal dalgacığın yumuşatılmış bu şekli ölçek fonksiyonu yardımıyla tespit edilir. Diğer taraftan ADD yönteminde dalgacık fonksiyonu, ölçek fonksiyonundan elde edilebilir. ADD için kullanılan ölçek $\varphi_{m,k}(n)$ ve dalgacık $\psi_{m,k}(n)$ fonksiyonları sırasıyla eşitlik (2.8) ve (2.9)’daki gibi tanımlanır.

$$\varphi_{m,k}(n) = 2^{-m/2} \varphi(2^{-m}n - k) \quad (2.8)$$

$$\psi_{m,k}(n) = 2^{-m/2} \psi(2^{-m}n - k) \quad n = 0, 1, \dots, M - 1 \quad (2.9)$$

Burada genişleme parametresi m (0,1,...,J-1) değer aralığında ve öteleme parametresi k (0,1,...,2^m - 1) değer aralığında yer alır. 2^{-m/2} ifadesi çarpım normalizasyonudur ve 2^J

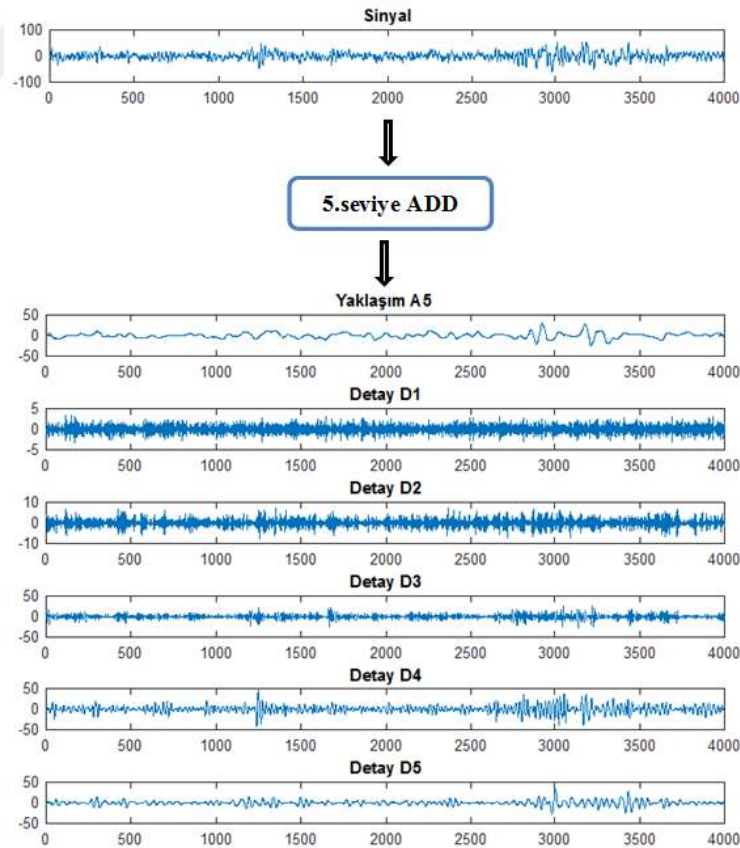
olarak seçilen M ifadesi sinyalin uzunluğu olmak üzere $J = \log_2(M)$ 'dir (Gonzalez ve Woods, 2002). Bir $x(n)$ sinyalinin m seviyesindeki ADD analizinde detay katsayıları dalgacık fonksiyonu yardımıyla eşitlik (2.10)'daki gibi ve yaklaşım katsayıları da ölçek fonksiyonu yardımıyla eşitlik (2.11)'deki gibi hesaplanır.

$$D_m(k) = \frac{1}{\sqrt{M}} \sum_{n=-\infty}^{\infty} x(n) 2^{-m/2} \psi(2^{-m}n - k) \quad (2.10)$$

$$A_m(k) = \frac{1}{\sqrt{M}} \sum_{n=-\infty}^{\infty} x(n) 2^{-m/2} \varphi(2^{-m}n - k) \quad (2.11)$$

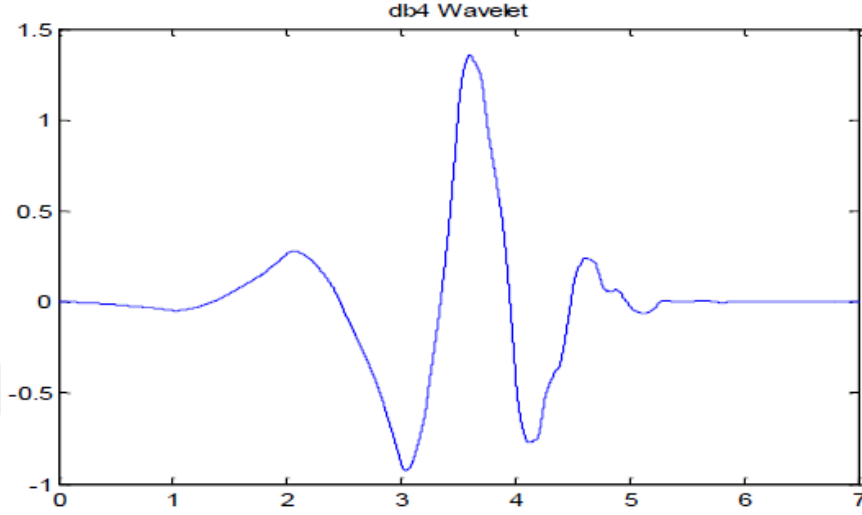
Burada D_m ve A_m ifadeleri m . dereceden detay ve yaklaşım katsayılarını temsil etmektedir (Gonzalez ve Woods, 2002).

Şekil 2.16'da bir sinyalin beşinci seviye ADD yöntemiyle alt bantlarına ayrıştırılması sonucunda elde edilen yaklaşım (A5) ve detay (D1-D5) alt bantları gösterilmektedir.



Şekil 2.16 Sinyalin beşinci seviye ADD yöntemiyle alt bantlarına ayrıştırılması

Dalgacık dönüşümünün en önemli bileşeni dalgacık fonksiyonudur. Filtre katsayılarının belirlenmesinde kullanılan bu fonksiyonun seçimi uygulanacak sinyale bağlıdır ve genellikle deneme yanılma yoluyla belirlenir. Şekil 2.17’de bu çalışmada kullanılan Daubechies 4 (db4) dalgacığı gösterilmiştir.



Şekil 2.17 Daubechies Dalgacıklarından db4 Dalgacığı
Misiti ve ark (2015)’ten modifiye edilmiştir.

2.6 Boyut İndirgeme Yöntemleri

2.6.1 Temel Bileşenler Analizi

1901 yılında Karl Pearson’un başlattığı TBA çalışmaları, 1933 yılında Hotelling tarafından geliştirilmiştir. TBA çok sayıda birbiri ile ilişkili değişkenler içeren veri setinin boyutlarını veri içerisinde var olan değişkenliğin olabildiğince korunarak daha az boyuta indirgenmesine yarayan bir dönüşüm tekniğidir. TBA, eldeki veriyi mümkün olduğunca az sayıda değişkenle temsil edebilecek en iyi dönüşümü belirlemeyi hedefler. Dönüşüm ile elde edilen değişkenler orijinal değişkenlerin temel bileşenleri olarak adlandırılır. İlk temel bileşen değişkenliği maksimum düzeyde açıklar ve diğer temel bileşenler değişkenliği açıklayıcılık bakımından azalacak şekilde sıralanır (Dunteman, 1989; Jolliffe, 2002).

p tane değişkenden oluşan bir veri seti p-boyutlu bir uzayda p sayıda orijinal eksenle veya p sayıda yeni eksenle temsil edilebilir. TBA'nın temel amacı orijinal eksenlerin oluşturduğu uzayı daha az boyutlu ve ortogonal olan yeni eksenler uzayına indirgemektir. Temel bileşenler veya yeni değişkenler, noktaların yeni eksenler üzerindeki izdüşümleridir. İlk yeni eksen ile gösterilen ilk temel bileşen toplam varyansı maksimum oranda açıklar. Daha sonra ilk eksenle ortogonal olan ikinci eksene karşılık gelen ikinci temel bileşen kalan varyansı maksimum ölçüde açıklar. Bu şekilde elde edilen temel bileşenler birbirleri ile ilişkisizdir ve orijinal değişken sayısını geçmeyecek şekilde istenilen sayıda temel bileşen elde edilebilir. Böylelikle az sayıda temel bileşenle toplam varyansın büyük bir kısmı açıklanabilir (Jolliffe, 2002; Masnan ve ark, 2012).

p sayıda değişken için elde edilebilecek temel bileşen sayısı en çok p tanedir ve temel bileşenler değişkenlerin lineer kombinasyonları şeklinde oluşturulurlar (Dunteman, 1989). Herhangi bir $\mathbf{x}$ rassal vektörünün doğrusal bir kombinasyonu

$$\mathbf{y}_1 = \sum_{k=1}^p a_{k1} x_k = \mathbf{a}_1' \mathbf{x} \quad (2.12)$$

şeklinde ifade edilsin. Burada $a_{11}, a_{21}, \dots, a_{p1}$; $\mathbf{a}_1$ vektörünün ağırlık katsayılarını, $\mathbf{y}_1$ de birinci temel bileşeni göstermektedir. $\mathbf{y}_1$ 'in varyansı $\mathbf{a}_1$ ağırlık vektörünün normuna ve yönüne bağlıdır. Ağırlık vektörünün normu büyüdükçe $\mathbf{y}_1$ 'in varyansı da büyüyecektir. Bu nedenle $\mathbf{a}_1$ 'in normu 1 olacak şekilde bir kısıt getirilerek maksimum varyansın elde edilmesi amaçlanır. $\|\mathbf{a}_1\| = 1$ olacak şekilde birinci temel bileşenin varyansı,

$$Var(\mathbf{a}_1) = E[\mathbf{y}_1^2] = \mathbf{a}_1' E[\mathbf{xx}'] \mathbf{a}_1 = \mathbf{a}_1' \mathbf{C}_x \mathbf{a}_1 \quad (2.13)$$

ifadesi ile elde edilir. Burada $\mathbf{C}_x = E[\mathbf{xx}']$ kovaryans matrisidir. $\mathbf{C}_x$ matrisinin öz değerleri $\lambda_1, \dots, \lambda_n$ ($\lambda_1 \geq \dots \geq \lambda_n$) ve bu öz değerlere karşılık gelen öz vektörleri $\mathbf{v}_1, \dots, \mathbf{v}_n$ 'in hesaplanmasıyla $Var(\mathbf{a}_1)$ ifadesini maksimize edecek sonuç ($\mathbf{a}_1 = \mathbf{v}_1$) elde edilecektir.

Birinci temel bileşen,

$$\mathbf{y}_1 = \mathbf{v}_1' \mathbf{x} \quad (2.14)$$

olarak ifade edilir. İkinci temel bileşen birinci temel bileşenle ilişkisiz olması kısıtı gözetilerek ($E[y_1 y_2] = 0$),

$$y_2 = v_2' x \quad (2.15)$$

olarak ifade edilir. Buna göre m. temel bileşen $1 \leq m \leq p$ için ve $E[y_k y_m] = 0$ ($k \neq m$) olmak üzere

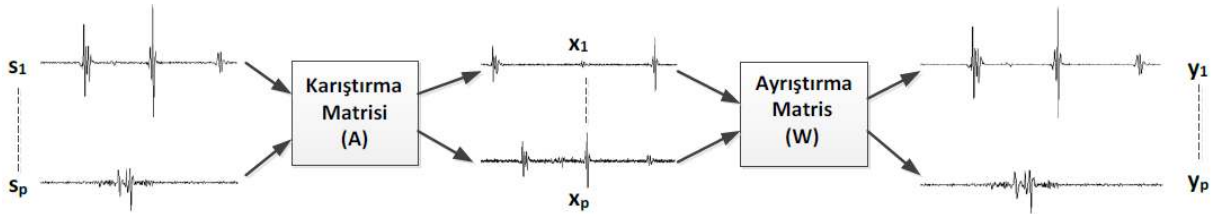
$$y_m = v_m' x \quad (2.16)$$

şeklinde ifade edilir (Dunteman, 1989).

2.6.2 Bağımsız Bileşen Analizi

Kör kaynak ayırıştırma yönteminin özel bir durumu olan BBA, rastgele değişkenler, ölçümler veya sinyallerin altında yatan gizli faktörlerin ortaya çıkarılmasında kullanılan istatistiksel bir hesaplama tekniğidir (Hyvärinen ve ark, 2001; Stone, 2004). BBA algoritması, biyomedikal sinyal işleme, ses sinyali ayırıştırma, telekomünikasyon, özellik çıkarma, görüntü işleme ve veri madenciliği gibi birçok alanda kullanılmaktadır.

Şekil 2.18'de BBA'nın modeli gösterilmiştir. Modelde p adet s_i orijinal sinyalleri, bilinmeyen bir karıştırma matrisi (**A**) ile karışarak p adet x_i karışım sinyalleri oluşuyor. BBA algoritmalarının amacı bu k adet karışımdan orijinal s_i sinyallerini bulabilecek bir ayırıştırma matrisini (**W**) bulmaya çalışmaktır. Sonuç olarak BBA algoritması ile s_i sinyallerinin tahminleri olan y_i sinyalleri elde edilir.



Şekil 2.18 BBA modeli

Bu işlem matematiksel olarak eşitlik (2.17) ile ifade edilir:

$$\begin{aligned} x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + \dots + a_{1p}s_p(t) \\ x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + \dots + a_{2p}s_p(t) \\ &\vdots \\ x_p(t) &= a_{p1}s_1(t) + a_{p2}s_2(t) + \dots + a_{pp}s_p(t) \end{aligned} \quad (2.17)$$

Denklem sisteminde $x_i(t)$, ölçüm sinyallerini, $s_i(t)$, orijinal kaynak sinyallerini ve a_{ij} ise karıştırma matrisinin bileşenlerini temsil etmektedir. Denklem sisteminde bilinen sadece x_i ölçüm sinyalleridir. Buna rağmen BBA, kaynakların bağımsız oldukları varsayımından yola çıkarak bu kaynakları ayırmayı başarmaktadır. Başka bir deyişle BBA, orijinal kaynaklar farklı olduğu için bu kaynakların istatistiksel anlamda bağımsız olduğunu varsayar. BBA bu varsayımın tersten ele alınarak işletilmesi mantığına dayanır. Yani kaydedilen karışım sinyallerinin istatistiksel olarak bağımsız olan sinyallere ayrıştırılmasıyla orijinal sinyaller tahmin edilir (Hyvärinen ve ark, 2001).

Denklem sisteminin matris notasyonu eşitlik (2.18)'deki gibidir.

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (2.18)$$

Bu eşitlikte $\mathbf{A}$ bilinmeyen bir karıştırma matrisidir. Elimizde sadece $x_i(t)$ ölçüm sinyalleri mevcutken eğer uygun bir tahmini ayrıştırma matrisi $\mathbf{W}$ bulunabilirse bu ayrıştırma matrisinin yardımıyla orijinal sinyallerin tahmini sinyallerini (bağımsız bileşenler, $y_i(t)$) bulmak mümkün olacaktır. Tahmin edilen ayrıştırma matrisi $\mathbf{W}$, $\mathbf{A}$ matrisinin tersine ne kadar yakın olursa bağımsız bileşenler orijinal sinyallere o kadar yakın olacaktır (Hyvärinen ve ark, 2001). Tahmin edilen sinyaller eşitlik (2.19) ile elde edilir.

$$\mathbf{Y} = \mathbf{W}\mathbf{X} \quad (2.19)$$

BBA algoritmalarının odaklandığı nokta en uygun karıştırma matrisinin tersini bulma yöntemidir. Ayrıştırma matrisi bulunurken bazı varsayımlar yapılmak zorundadır. BBA ile $\mathbf{W}$ matrisi bulunurken yapılan varsayımlardan birincisi karışımı oluşturan orijinal sinyallerin istatistiksel olarak birbirinden bağımsız olması ve ikincisi ise karışımı oluşturan sinyallerden sadece biri normal dağılıma sahip olabilirken diğerlerinin normal dağılıma sahip olmamasıdır. Bu varsayımlardan sonra BBA'nın başarısı için bir diğer varsayım ise en az orijinal sinyal sayısı kadar karışım sinyalinin mevcut olmasıdır. Diğer ayrıştırma algoritmaları arasında

BBA'yı güçlü kılan özellik, sağlanması gereken varsayımların az olmasıdır. BBA'nın dezavantajı ise karışımı oluşturan kaynakların genliklerini ve karışım sıralarını doğru olarak tahmin edememesidir (Hyvärinen ve ark, 2001; Stone, 2004).

BBA algoritmaları uygulanmadan önce bir takım ön işlemler gerçekleştirilir. İlk olarak $\mathbf{X}$ veri matrisinin merkezlenmesidir. Bunun anlamı, tüm değerlerden ortalamanın çıkartılmasıdır. Böylece $\mathbf{X}$ veri matrisi sıfır ortalamalı hale getirilmiş olur. Bu işlem BBA algoritmasını kolaylaştırır. Daha sonra yapılacak ön işlem *beyazlatma*'dır. Beyazlatma, karışım sinyallerinin ilişkisiz ve birim varyansa sahip olmaları için yapılan bir dönüşümdür. Başka bir ifade ile herhangi bir $\mathbf{x}$ vektörü beyaz ise $\mathbf{x}$ vektörü için aşağıdaki eşitlikler sağlanır:

$$\mathbf{m}_x = 0$$

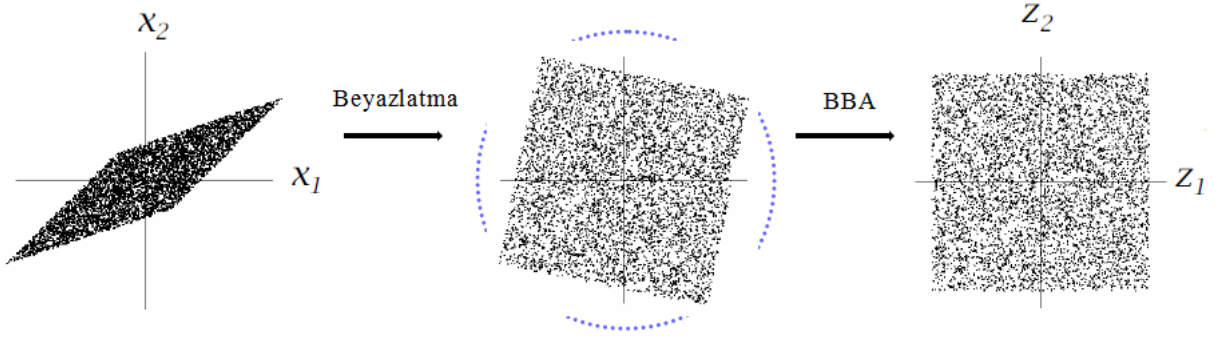
ve

$$\mathbf{C}_x = \mathbf{R}_x = \mathbf{I}$$

Burada $\mathbf{m}_x$, $\mathbf{x}$ vektörünün ortalamasını, $\mathbf{C}_x$, kovaryans matrisini, $\mathbf{R}_x$, korelasyon matrisini ve $\mathbf{I}$, birim matrisi temsil etmektedir. Ayrıca $\mathbf{x}$ beyaz bir vektör ve $\mathbf{V}$ ortogonal bir matris olmak üzere $\mathbf{z} = \mathbf{V}\mathbf{x}$ dönüşümü sonucu elde edilen $\mathbf{z}$ vektörü de beyazdır. Yani ortogonal dönüşümlerde beyazlık özelliği korunmaktadır (Hyvärinen ve ark, 2001).

BBA'da beyazlatma işlemi için TBA yaygın olarak kullanılmaktadır. TBA'da elde edilecek bileşen sayısı yapılan çalışmanın amacına ve açıklanacak toplam varyans miktarına göre araştırmacı tarafından belirlenebileceği için beyazlatma işlemi yanında boyut indirgeme işlemi de yapılabilir. Karışım sinyalleri ilişkisiz hale getirildikten sonra eşit uzunlukta olacak şekilde normalize edilir. Böylelikle beyazlatma işlemi tamamlanmış olur. Son olarak BBA uygulanarak orijinal sinyaller tahmin edilir (Hyvärinen ve ark, 2001; Stone, 2004).

Şekil 2.19'da karışmış durumdaki iki bağımsız sinyal için yapılan beyazlatma ve BBA sonucu yapılan dönüşümler görsel olarak verilmiştir. Burada ilişkili durumdaki karışmış iki bağımsız sinyal beyazlatma işlemi ile önce ilişkisiz sinyallere daha sonra uygulanan BBA ile orijinal sinyallerin tahminleri olan bağımsız bileşenlere dönüştürülmüştür.



Şekil 2.19 Beyazlatma ve BBA dönüşümü

Bağımsız bileşenleri tahmin etmenin başlangıç adımı, ölçümlerin ilişkisiz ve birim varyansa sahip olmaları için yapılan bir dönüşüm olan beyazlatmadır (Hyvärinen ve ark., 2001). Beyazlatmadan sonra $\mathbf{W}$ ayrıştırma matrisi tahmin edilir. BBA ile $\mathbf{W}$ matrisi tahmin edilirken hedeflenen, tahmini sinyallerin olasılık yoğunluk dağılımını Gauss Dağılımı'ndan en çok uzaklaştıran $\mathbf{W}$ matrisini bulmaktır. Bu kriterin dayanak noktası ise merkezi limit teoreminden gelir. Merkezi limit teoremine göre, farklı dağılımlara sahip bağımsız rastgele değişkenlerin toplamından elde edilen sinyalin olasılık dağılımı her zaman değişkenlerin tek başına olduğundan daha fazla Gauss Dağılımı'na benzeyecektir. Başka bir deyişle rasgele değişken sayısı sonsuza gittikçe elde edilen toplamsal sinyalin dağılımı da Gauss Dağılımı'na yakınsayacaktır (Girolami, 1999). Bu bilgi karışım halindeki sinyallerden orijinal sinyalleri elde etmek için kullanılır. Karışım halindeki sinyalin dağılımından daha az Gauss Dağılımı'na sahip bir sinyal bulunursa bu sinyalin bağımsız bileşen olma ihtimali vardır, denilir.

Karıştırma matrisi $\mathbf{W}$ 'nin tahmini için bir takım algoritmalar geliştirilmiştir. Bu algoritmalarından birisi de Hyvärinen (1999) tarafından geliştirilen Hızlı Sabit-Nokta Algoritması'dır ve kısaca FastICA olarak adlandırılır. FastICA, hızlı yakınsama sağlaması, kolay uygulanabilir ve güvenilir sonuçlar vermesi nedeniyle en çok kullanılan BBA algoritmasıdır. FastICA'nın birden fazla bağımsız bileşeni teker teker tahmin eden algoritmasının adımları aşağıdaki gibidir:

1. Veri seti ortalaması sıfır olacak şekilde merkezileştirilir.
2. Veri seti beyazlatılarak $\mathbf{Z}$ elde edilir ($\mathbf{Z} = \mathbf{VX}$).
3. Tahmin edilecek bağımsız bileşen sayısı olan m belirlenir, $p \leftarrow 1$ alınır.
4. $\mathbf{w}_p$ 'nin başlangıç değeri rassal olarak belirlenir.

5. g için $g(y) = \tanh(a_1 y)$, $g(y) = ye^{(-y^2/2)}$ ya da $g(y) = y^3$ fonksiyonlarından birisi kullanarak $\mathbf{w}_p \leftarrow E\{\mathbf{z}g(\mathbf{w}_p^T \mathbf{z})\} - E\{g'(\mathbf{w}_p^T \mathbf{z})\}\mathbf{w}$ güncellenir.
6. $\mathbf{w}_p \leftarrow \mathbf{w}_p - \sum_{j=1}^{p-1} (\mathbf{w}_p^T \mathbf{w}_j) \mathbf{w}_j$ ortogonalleştirilmesi yapılır.
7. $\mathbf{w}_p \leftarrow \mathbf{w}_p / \|\mathbf{w}_p\|$ standartlaştırması yapılır.
8. $\mathbf{w}_p$ yakınsamamışsa 5. adıma geri dönülür.
9. $p \leftarrow p + 1$ alınır, $p \leq m$ ise 4. adıma geri dönülür.

Algoritmada kullanılan g fonksiyonunun seçimi önem taşımaktadır (Gadhok ve Kinsner, 2004). Yapılan çalışmalara dayanarak sapan değerlere karşı $g(y) = ye^{(-y^2/2)}$ Gauss fonksiyonunun kullanımı önerilmektedir (Hyvärinen, 1999; Gadhok ve Kinsner, 2004).

2.7 Sınıflandırma Yöntemleri

Sınıflandırma, verileri ayırmak ve ait oldukları sınıflara atamak amacıyla tahmin modelleri oluşturan istatistiksel bir metottur. Sınıflandırma işleminde oluşacak hatayı (yanlış sınıflandırma) minimum yapacak olan sınıflandırma yönteminin seçilmesi gerekmektedir.

Bu çalışmada sınıflandırma için LDA, lineer ve radyal tabanlı kernel fonksiyonlarının kullanıldığı DVM yöntemleri kullanıldı.

2.7.1 Diskriminant Analizi

Diskriminant Analizi (DA), X veri setindeki değişkenlerin iki veya daha fazla gruba ayrılmasını sağlayan, birimlerin p tane özelliğini ele alarak bu birimlerin gerçek gruplarına, sınıflarına optimal düzeyde atanmalarını sağlayacak fonksiyonlar türeten bir yöntemdir (Hardle ve Simar, 2003; Özdamar, 2013).

DA aracılığıyla elde edilen diskriminant (ayırma) fonksiyonları, tahmin değişkenlerinin doğrusal bileşimlerinden oluşur. Ayırma fonksiyonları gruplar arası farklılığa etki eden tahmin değişkenlerinin hangileri olduğunu ortaya çıkarır. Gruplar arası farklılığa etki eden bu değişkenlere de diskriminant (ayırma) değişkenleri adı verilir. Diskriminant analizinin bir diğer işlevi ise gruplardan herhangi birine ait olan fakat hangi gruptan geldiği bilinmeyen bir birimin ait olduğu gruba en az hata ile atanmasını sağlamaktır (Johnson ve Wichern, 2007).

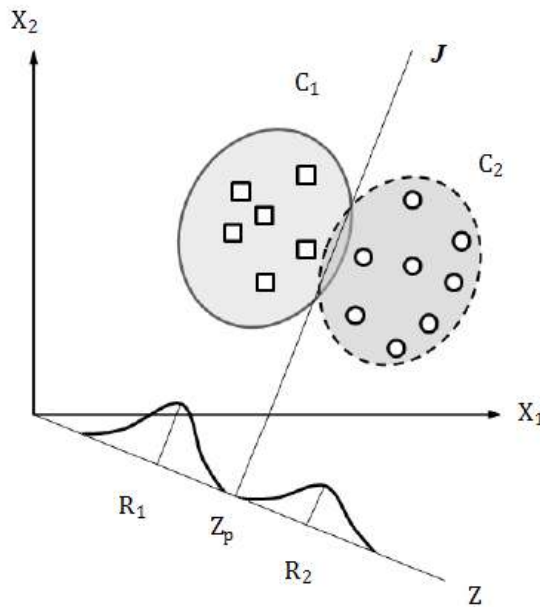
DA yapılabilmesi için birtakım varsayımlarının sağlanması gereklidir. Bu varsayımlar:

- X veri matrisi çok değişkenli normal dağılım göstermelidir.
- Grup kovaryans matrisleri homojen olmalıdır ve X matrisinde yer alan değişkenler ortak kovaryans matrisine sahip çok değişkenli normal dağılım gösteren toplumdaki çekilmiş olmalıdır.
- Değişkenlerin ortalamaları ve varyansları arasında bir korelasyon bulunmamalıdır.
- Değişkenler arasında çoklu bağımlılık bulunmamalıdır.
- Değişkenler arasında çok yüksek korelasyon ($r > 0.80$) bulunmamalıdır.

X matrisi gereğinden fazla ve gereksiz değişken içermemeli, grupların birbirlerinden ayrılmasını sağlayacak kadar doğru ve gerekli değişkenleri içermelidir (Özdamar, 2013).

2.7.1.1 Diskriminant Analizine Geometrik Yaklaşım

Diskriminant analizinin geometrik yorumunu yapmak için Şekil 2.20’de iki gruplu diskriminant analizinin grafik gösterimi verilmiştir.



Şekil 2.20 İki gruplu diskriminant analizinin grafik gösterimi

Şekil 2.20’de C_1 ve C_2 sırasıyla n_1 ve n_2 sayıda gözlemden oluşan iki farklı sınıf, X_1 ve X_2 bu sınıfları ayırmada kullanılacak olan değişkenleri göstermektedir. Şekilde “□” C_1 grubunun ve “○” C_2 grubunun birimlerini göstermektedir. Elipsler ise ilgili gruba ait birimlerin yaklaşık %95 veya daha fazlasını içermektedir. İki elipsin kesiştiği noktaları

birleştiren bir J doğrusu çizilir. Bu doğru yeni bir Z eksenine üzerine dik olarak uzatılırsa R_1 ve R_2 tek değişkenli dağılımlarının kesişim kümesi diğer çizilebilecek doğrulara kıyasla daha küçük olur. Böylece iki grubun ayrımı en iyi şekilde sağlanmış olur (Sharma, 1995).

Z eksenine diskriminant fonksiyonunu temsil etmektedir. Z eksenine ve elipslerin kesişim noktalarını birleştiren doğrunun kesiştiği noktaya (Z_p) kritik değer denilmektedir. X_1 ve X_2 değişkenleriyle tanımlanan gözlemler, diskriminant fonksiyonu yardımıyla bu değişkenlerin lineer bileşimi olarak tek bir diskriminant değerine dönüştürülmektedir. Böylece her gözlem Z ekseninde bir nokta olarak gösterilebilir. Gözlemlerin Z eksenindeki izdüşümleri kritik noktanın sağında veya solunda olmasına göre noktanın temsil ettiği gözlem C_1 veya C_2 grubuna atanmaktadır. Eğer grupları ayırmak için kullanılan değişken sayısı ya da grup sayısı ikiden fazlaysa grafikte gösterme olanağı güçleşmektedir (Sharma, 1995).

2.7.1.2 İki Grup için Lineer Diskriminant Analizi

$A(\pi_1)$ ve $B(\pi_2)$ isimli iki toplum olsun. Bu toplumlardan n birimlik p tane birbirleri ile ilişkili gözlem yapılsın. X veri matrisinde gözlemler toplumdan topluma az ya da çok farklılıklar gösterir. π_1 toplumundan alınan örnekler için X_1 gözlem matrisi ve π_2 toplumundan alınan örnekler için X_2 gözlem matrisi elde edilir. Çok değişkenli X gözlem matrisi π_1 ve π_2 toplumlarına ilişkin gözlemleri içeren X_1 ve X_2 gözlem matrislerinin birleştirilmesiyle elde edilir (Özdamar, 2013).

g_1 ve g_2 sırasıyla π_1 ve π_2 toplumlarından n_1 ve n_2 hacimli gruplar olsun. Bu gruplara ilişkin kovaryans matrisleri S_1 ve S_2 olmak üzere toplumların aynı kovaryans matrisine (Σ) sahip oldukları kabul edilerek ($S_1=S_2$), g_1 ve g_2 gruplarının ortak kovaryans matrisi S , eşitlik (2.20)'deki gibi tanımlanır.

$$S = S_{\text{ortak}} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{(n_1 + n_2 - 2)} \quad (2.20)$$

X gözlem matrisi kullanılarak elde edilen çok değişkenli gözlem matrisi tek değişkenli Y değerlerine dönüştürülür. Bu Y değerleri X gözlem matrisinin doğrusal bileşenleridir.

LDA ile her bir grup için birer ayırma fonksiyonu aşağıdaki gibi hesaplanır. İki grup arasındaki p değişken için oluşturulan doğrusal ayırma fonksiyonu aşağıda verilmiştir:

$$Y_i = b_{i0} + b_{i1}X_1 + b_{i2}X_2 + \dots + b_{ip}X_p \quad i: 1,2 \text{ grup sayısı} \quad (2.21)$$

Bu fonksiyonda b_{i0} sabit değeri, b_{ij} ise doğrusal bileşenleri belirtmektedir. Doğrusal bileşenlere kanonik değişkenler adı da verilmektedir. Her bir grup için değişkenlere ilişkin doğrusal bileşenler, değişkenlerin ayırma fonksiyonundaki etkinliklerini ve belirleyiciliklerini göstermektedir (Özdamar, 2013).

Doğrusal bileşenler (2.22)'te verilen eşitlik ile hesaplanır.

$$b_{ij} = \mathbf{S}^{-1}(\bar{\mathbf{x}}_i) \quad i = 1,2; j = 1, \dots, p \quad (2.22)$$

Burada $\mathbf{S}^{-1}$, ortak kovaryans matrisinin tersini ve $\bar{\mathbf{x}}_i$ ise i.grup için ortalama vektörünü temsil etmektedir. b_{ij} katsayılarına grup ayırma fonksiyonu katsayıları ya da kanonik değişkenler adı verilir. Bazen katsayılar ölçeklendirilerek ya da normalize edilerek kullanılmaktadır. Normalizasyonun temel amacı katsayıları genel katsayılar içinde ağırlıklandırarak elemanların kolay yorumlanmasını sağlamaktır. İki tür normalizasyon yaklaşımı vardır.

$$1 - b^* = \frac{b}{\sqrt{b'b}} \quad \text{veya} \quad b^* = \frac{b}{b_1}$$

Burada b_1 en büyük doğrusal bileşendir.

Her iki normalizasyon yaklaşımında da katsayılardan birinin diğerlerine göre önemli olarak ağırlığını artırmak ve diğerlerine göre daha önemli değişkenleri modelde etkin olarak görmek amaçlanmaktadır (Sharma, 1995; Özdamar, 2013).

Gruplar arasındaki farklılığı maksimize edecek bir ayırma fonksiyonu aracılığı ile grupları birbirinden ayırmak mümkün olacaktır. Bu nedenle ortak bir ayırma fonksiyonu belirlenir. i ve j grupları arasındaki ayırma fonksiyonu;

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p \quad (2.23)$$

şeklinde yazılır. Bu fonksiyondaki b_i doğrusal bileşenleri ortalama fark vektörü aracılığı ile aşağıdaki gibi bulunur.

$$b_i = \mathbf{S}^{-1}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) \quad i = 1, 2, \dots, g \text{ grup sayısı} \quad (2.24)$$

Sabit değer b_0 katsayısı ise;

$$b_0 = -(1/2)\bar{\mathbf{x}}'\mathbf{S}^{-1}\bar{\mathbf{x}} \quad (2.25)$$

biçiminde hesaplanır.

$\mathbf{S}$ ortak kovaryans matrisi ve $\bar{\mathbf{x}}_i$ i. grup ortalama vektörü olmak üzere her bir grubun b_i katsayılar vektörü aşağıdaki gibi hesaplanır.

$$b_{ij} = \mathbf{S}^{-1}\bar{\mathbf{x}}_i \quad i = 1, 2; j = 1, \dots, p \quad (2.26)$$

Sabit değer ise;

$$b_{i0} = -(1/2)\bar{\mathbf{x}}_i'\mathbf{S}^{-1}\bar{\mathbf{x}}_i \quad (2.27)$$

biçiminde hesaplanır.

Gruplara göre belirlenen sabit ve kanonik katsayılar değişken değerleri ile çarpılarak doğrusal ayırma fonksiyonları hesaplanır.

1.grup için ayırma fonksiyonu;

$$Y_1 = b_{10} + b_{11}X_{11} + b_{12}X_{12} + \dots + b_{1p}X_{1p} \text{ eşitliğinde değerler yerine konularak belirlenir.}$$

2.grup için ayırma fonksiyonu;

$$Y_2 = b_{20} + b_{21}X_{21} + b_{22}X_{22} + \dots + b_{2p}X_{2p} \text{ eşitliğinde değerler yerine konularak belirlenir.}$$

g.grup için ayırma fonksiyonu;

$$Y_g = b_{g0} + b_{g1}X_{g1} + b_{g2}X_{g2} + \dots + b_{gp}X_{gp} \text{ eşitliğinde değerler yerine konularak belirlenir.}$$

p değişkenli g grup arasındaki karesel uzaklığı veren ve Mahalanobis tarafından ileri sürülen D^2 uzaklığı aşağıdaki gibi hesaplanır.

$$D_{ij}^2 = (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)'\mathbf{S}^{-1}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) \quad (2.28)$$

D^2 uzaklığının i ve j gruplarını birbirinden ayırmada etkin rol oynayıp oynamadığı test edilebilir. Bu amaçla Hotelling T^2 yaklaşımı ile D^2 'nin önemliliği test edilebilir.

$$T^2 = \left(\frac{n_1 * n_2}{n_1 + n_2} \right) D^2 \quad (2.29)$$

T^2 'nin önemliliğinin belirlenmesi için F yaklaşımından yararlanılır.

$$F = \frac{(n_1 + n_2 - p - 1)}{p(n_1 + n_2 - 2)} T^2 \quad (2.30)$$

F'nin önemliliği, p, (n_1+n_2-p-1) serbestlik dereceli F dağılımının kritik değerleri kullanılarak belirlenir (Sharma, 1995; Özdamar, 2013).

Bağımlı değişken iki kategorili olduğu zaman yalnızca bir tane diskriminant fonksiyonu hesaplanır. Bu fonksiyonun ne kadar önemli olduğunu belirlemek için kanonik korelasyon, öz değer ve Wilks' Lamda değerlerine bakılır. Kanonik korelasyon, diskriminant skorları ve gruplar arasındaki ilişkiyi ölçer ve açıklanan toplam varyansı gösterir. Öz değer istatistiği ise ne kadar büyük olursa bağımlı değişkendeki varyansın o kadar büyük bir kısmı diskriminant fonksiyonu ile açıklanır. Kesin bir değer olmamakla birlikte 0,40'tan büyük olan öz değerler, iyi olarak kabul edilir. Wilks' Lamda istatistiği, diskriminant skorlarındaki toplam varyansın gruplar arasındaki farklar tarafından açıklanmayan kısmını gösterir (Kalaycı, 2006).

2.7.1.3 Lineer Diskriminant Analizinde Gözlemlerin Sınıflandırılması

LDA ile yanlış sınıflandırılmış gözlemleri belirlemek ve yanlış sınıflandırma olasılıklarını hesaplamak mümkündür.

İki toplumdan oluşan bir gözlem skalasında, R_1 bölgesi π_1 toplumunda birimlerin gözlenme alanı ve R_2 bölgesi π_2 toplumunda gözlenme alanı, π_1 toplumu gözlenme olasılığı p_1 ve π_2 toplumu gözlenme olasılığı p_2 olmak üzere bir birimin gerçekte π_1 toplumuna ait olduğu halde π_2 olarak ayrılmasının koşullu olasılığı;

$$P(2|1) = P(X \in R_2 | \pi_1) \quad (2.31)$$

şeklinde hesaplanır. Aynı şekilde bir birimin gerçekte π_2 toplumuna ait olduğu halde π_1 olarak ayrılmasının koşullu olasılığı;

$$P(1|2) = P(X \in R_1|\pi_2) \quad (2.32)$$

şeklinde hesaplanır (Rencher, 2002; Johnson ve Wichern, 2007).

Toplamları ayırmada ön olasılık (prior probability) değerleri biliniyorsa yukarıdaki eşitlikler aşağıdaki gibi ifade edilirler:

$$P(1|1) = P(X \in R_1|\pi_1)p_1$$

$$P(2|2) = P(X \in R_2|\pi_2)p_2$$

$$P(2|1) = P(X \in R_2|\pi_1)p_1$$

$$P(1|2) = P(X \in R_1|\pi_2)p_2$$

Burada $p_1+p_2=1$ olmak zorundadır. Bu olasılıklar yardımı ile yanlış sınıflandırma maliyeti ($c(i|j)$) matrisi Tablo 2.2'deki gibi oluşturulur.

Tablo 2.2 Yanlış sınıflandırma maliyeti

		Gerçek Toplum	
		π_1	π_2
Sınıflandırma	π_1	0	$c(1 2)$
	π_2	$c(2 1)$	0

Birimleri yanlış ayırmanın ya da sınıflamanın maliyeti ise $c(1|2)$ ve $c(2|1)$ şeklinde gösterilir. Buradan Beklenen Yanlış Sınıflandırma Maliyeti (BYSM) ise

$$BYSM = c(2|1)P(2|1)p_1 + c(1|2)P(1|2)p_2 \quad (2.33)$$

şeklinde hesaplanır (Özdamar, 2013).

Diskriminant analizinde sınıflandırma kuralının temelinde BYSM değerinin minimize edilmesi önemlidir. BYSM'yi minimize eden R_1 ve R_2 bölgeleri aşağıdaki gibi ifade edilir:

$$R_1: \frac{f_1(x)}{f_2(x)} \geq \left(\frac{c(1|2)}{c(2|1)} \right) \left(\frac{p_2}{p_1} \right) \quad (2.34)$$

$$R_2: \frac{f_1(x)}{f_2(x)} < \left(\frac{c(1|2)}{c(2|1)} \right) \left(\frac{p_2}{p_1} \right) \quad (2.35)$$

Burada $f_1(x)$ ve $f_2(x)$ sırasıyla π_1 ve π_2 'ye ait yoğunluk fonksiyonlarını temsil etmektedir (Özdamar, 2013).

Yeni gözlenen ve $\mathbf{x}_0$ gözlem vektörüne sahip bir birimin son sınıflandırma olasılıklarına göre hangi sınıfa atanacağı belirlenir ve son sınıflandırma olasılıkları aşağıdaki gibi hesaplanır:

$$P(\pi_1|\mathbf{x}_0) = \frac{p_1 f_1(\mathbf{x}_0)}{p_1 f_1(\mathbf{x}_0) + p_2 f_2(\mathbf{x}_0)} \quad (2.36)$$

$$P(\pi_2|\mathbf{x}_0) = \frac{p_2 f_2(\mathbf{x}_0)}{p_1 f_1(\mathbf{x}_0) + p_2 f_2(\mathbf{x}_0)} \quad (2.37)$$

İki grup için minimum BYSM'ye sahip bölgeler;

$$R_1: \exp \left[-\frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}}_1)' \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_1) + \frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}}_2)' \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_2) \right] \geq \left(\frac{c(1|2)}{c(2|1)} \right) \left(\frac{p_2}{p_1} \right) \quad (2.38)$$

$$R_1: \exp \left[-\frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}}_1)' \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_1) + \frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}}_2)' \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_2) \right] < \left(\frac{c(1|2)}{c(2|1)} \right) \left(\frac{p_2}{p_1} \right) \quad (2.39)$$

biçiminde belirlenir ve $\mathbf{x}_0$ gözleminin π_1 veya π_2 'den hangisine atanacağı aşağıdaki kurala göre belirlenir:

Eğer;

$$\left[(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{S}^{-1} \mathbf{x}_0 - \frac{1}{2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{S}^{-1} (\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2) \right] \geq \ln \left(\frac{c(1|2)}{c(2|1)} \right) \left(\frac{p_2}{p_1} \right) \quad (2.40)$$

ise $\mathbf{x}_0$ gözlemi π_1 toplumuna aittir. Diğer durumda π_2 toplumuna aittir (Özdamar, 2013).

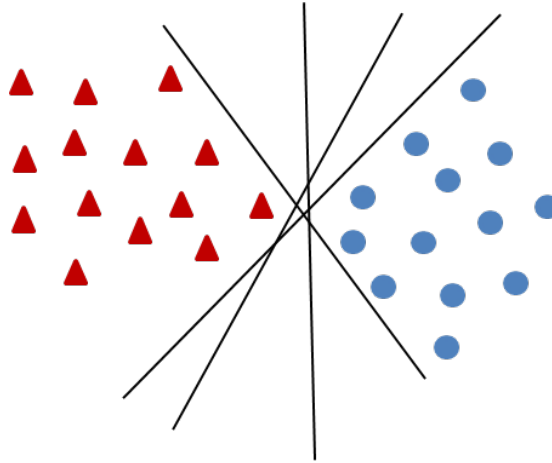
2.7.2 Destek Vektör Makinesi

DVM algoritması, örüntü tanıma, karar verme ve sınıflandırma için güçlü bir makine öğrenme yöntemidir. Özellikle EEG sinyallerinin sınıflandırılmasında sıklıkla kullanılmakta ve başarılı sonuçlar vermektedir (He ve ark, 2004, Lima ve ark, 2010; Subaşı ve Gürsoy, 2010).

DVM istatistiksel öğrenme teorisi üzerine inşa edilmiştir. Sinir ağları ve karar ağaçları gibi sadece gözlemsel sınıflandırma hatasını minimize etmeye çalışan istatistiksel öğrenme metotlarının aksine DVM gözlemsel sınıflandırma hatasını ve sınıflandırmadaki geometrik uzaklığı eş zamanlı olarak minimize eden bir yöntemdir (Burges, 1998; Sanei ve Chambers, 2007).

DVM eğitim aşamasında, iki gruptaki gözlemler arası uzaklık maksimum olacak şekilde, en iyi ayırıcı hiper düzlemi bulmaya çalışır. Bu aşamada esas olan, ayırıcı hiper düzlemi orijinal öznitelik uzayında değil orijinal öznitelik uzayının lineer olmayan bir dönüşüm fonksiyonu yardımıyla eşleştirildiği yüksek boyutlu bir uzay içerisinde bulmaktır (Burges, 1998; Sanei ve Chambers, 2007).

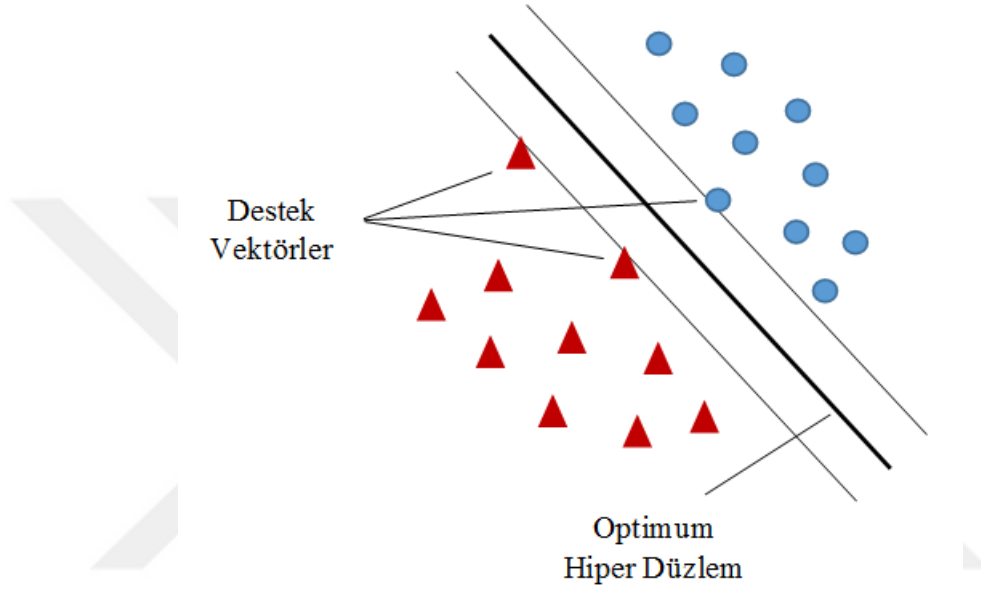
İki farklı sınıfı birbirinden ayırmak için sonsuz sayıda hiper düzlem çizilebilir (Şekil 2.21).



Şekil 2.21 Ayırıcı hiper düzlemler

Şekil 2.21’de verilen sınıflar arasında sonsuz sayıda hiper düzlem çizilebilir. DVM, hata toleransını minimuma indirmek için çizilebilecek hiper düzlemler arasında her iki sınıfa da maksimum uzaklıkta olacak şekilde çizilen optimum hiper düzlemi bulmayı amaçlar (Cortes ve Vapnik, 1995; Burges, 1998; Sanei ve Chambers, 2007).

DVM ile sınıflandırma mantığının temeli her iki sınıfın uç noktalarında bulunan ve eğitim verilerinin arasından seçilen destek vektörleridir. Ayrıca düzlemin optimum olmasıyla genelleme yeteneği de en üst düzeyde olacaktır. Optimum hiper düzlem, iki sınıfın destek vektörleri arasındaki uzaklığın maksimum olduğu durumu sağlayan ayırıcı hiper düzlemdir (Burges, 1998; Sanei ve Chambers, 2007). Şekil 2.22’de gösterildiği gibi optimum hiper düzlem bahsedilen aralığın tam ortasından geçmelidir.



Şekil 2.22 Destek vektörleri ve optimum hiper düzlem

Şekil 2.22’de basit bir DVM sınıflandırıcısının yapısı gösterilmektedir. Burada amaç optimum hiper düzlemin bulunmasıdır. Giriş verileri $\{-1, +1\}$ değerini alan iki sınıfa ayrılırlar. Destek vektörleri bulundukları sınıfın optimum hiper düzleme en yakın elemanlarıdır ve hiper düzlemin optimum olması için destek vektörleriyle optimum hiper düzlem arasındaki uzaklığın maksimum olması gerekir (Burges, 1998; Sanei ve Chambers, 2007).

DVM ile yapılacak sınıflandırmalarda ortaya çıkabilecek iki durum vardır: Verilerin lineer şekilde ayrılabilen veya ayrılamayan bir yapıda olmasıdır.

Gerçek hayatta karşılaşılabilecek problemlerin büyük çoğunluğu çok sayıda farklı bileşenden oluşur ve lineer olmayan bir yapıdadır. Lineer olarak ayrılabilen veriler arasındaki maksimum uzaklığın bulunması zor değildir, fakat lineer şekilde ayrılamayan veriler öncelikle lineer şekilde ayrılacakları başka bir uzaya aktarılmalıdırlar (Sanei ve Chambers, 2007).

2.7.2.1 Lineer Olarak Ayrılabilme Durumu için DVM

Eğitim için kullanılacak N sayıda elemandan oluşan verinin $X = \{x_i, y_i\}$, $i = 1, 2, \dots, N$ olduğunu varsayalım. Burada $y_i \in \{-1, +1\}$ sınıf etiketi değerleri ve $x_i \in R^d$ özellikler vektörüdür. Lineer olarak ayrılabilme durumunda, iki değerli bu veriler doğrudan bir hiper düzlem ile ayrılacaktır. DVM'lerin amacı bu hiper düzlemin iki ayrı sınıfta bulunan örnek grubuna eş uzaklıkta olmasını sağlamaktır (Cortes ve Vapnik, 1995; Burges, 1998).

DVM ilk işlem olarak lineer şekilde ayrılamayan verileri yüksek boyutlu bir uzaya taşır. Bu uzayın boyutu, bağımsız değişken sayısından büyük olur. Bu işlemin yapılmasının amacı verilerin bir hiper düzlem ile ayrılmasını sağlamaktır. x_i hiper düzlem üzerinde herhangi bir nokta, w hiper düzlemin normali ve $|b|/\|w\|$ hiper düzlemin orijine dik uzaklığı olmak üzere, DVM algoritması,

$$w' \cdot x_i + b = 0 \quad (2.41)$$

koşulunu sağlayan optimum hiper düzlemi bulmaya çalışır (Cortes ve Vapnik, 1995; Burges, 1998). Bunun için eğitim veri setinin eşitsizlik (2.42) ve (2.43)'ü sağlaması gerekmektedir:

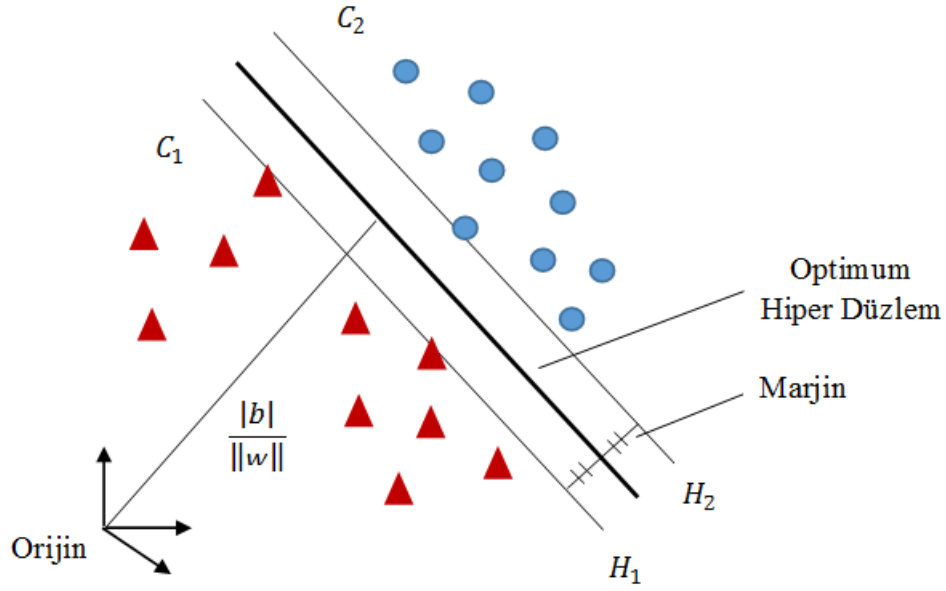
$$w' \cdot x_i + b \geq +1, \quad y_i = +1 \text{ için} \quad (2.42)$$

$$w' \cdot x_i + b \leq -1, \quad y_i = -1 \text{ için} \quad (2.43)$$

Bu iki eşitsizlik tek bir eşitsizlik olarak,

$$y_i(w' \cdot x_i + b) - 1 \geq 0 \quad (2.44)$$

şeklinde ifade edilir (Cortes ve Vapnik, 1995). Eğitim veri seti için olası tüm hiper düzlemler bu son eşitsizlik ile temsil edilebilir. Bu eşitsizliği sağlayan optimum hiper düzlemin iki tarafındaki en yakın noktalara olan dik uzaklıkları toplamı marjın olarak adlandırılır ve optimum hiper düzlem için marjın maksimum olur. Optimum hiper düzlemi bulmak için uygun w ve b değerleri hesaplanır. Şekil 2.23'te lineer ayrılabilme durumuna ilişkin temsili gösterim verilmiştir.



Şekil 2.23 Lineer ayrılabilirlik durumu

Bağıntı (2.42)'nin eşitlik durumu ($w' \cdot x_i + b = +1$) Şekil 2.23'teki C_2 sınıfını ayıran H_2 ayırıcı düzlemine, bağıntı (2.43)'ün eşitlik durumu ($w' \cdot x_i + b = -1$) ise C_1 sınıfını ayıran H_1 ayırıcı düzlemine karşılık geldiğini varsayalım. H_1 hiper düzleminin orijine dik uzaklığı $|1 - b|/\|w\|$ ve H_2 hiper düzleminin orijine dik uzaklığı $|-1 - b|/\|w\|$ olur. H_1 ve H_2 hiper düzlemlerinin optimum hiper düzleme ($w' \cdot x_i + b = 0$) uzaklıkları ise $1/\|w\|$ olur. Bu iki hiper düzlem veya başka bir deyişle eğitim verisine ait iki veri seti arasındaki uzaklık $2/\|w\|$ olur. Burada H_1 ve H_2 hiper düzlemlerinin aynı normale sahip olduklarına (paralel olduklarına) ve bu düzlemler arasında eğitim veri setine ait hiçbir gözlem bulunmadığına dikkat edilmelidir. H_1 ve H_2 hiper düzlemleri arasındaki uzaklık ($2/\|w\|$) marjine eşittir. DVM algoritmasıyla yapılmaya çalışılan, marjini maksimize etmektir. Dolayısıyla, $\|w\|$ 'nin minimum olmasını sağlamaktır. H_1 ve H_2 hiper düzlemlerinin üzerinde yer alan gözlemler ise destek vektörleri olarak adlandırılır (Burges, 1998).

Maksimum sınırın bulunması işlemi;

$$\min \left(\frac{1}{2} \|w\|^2 \right) \quad (2.45)$$

olarak elde edilir. Buna bağlı sınırlamalar ise;

$$y_i(w' \cdot x_i + b) - 1 \geq 0, \quad i = 1, \dots, k$$

şeklinde ifade edilir. Bu ifadeler problemin ikinci dereceden bir optimizasyon problemi olduğunu gösterir. Ancak çok yüksek boyutlu uzaylar için problemin çözümü kolay değildir. Bu nedenle problem pratik olarak çözülebileceği ikincil forma dönüştürülür (Burges, 1998). Problemin ikincil forma dönüştürülmesi iki adımda gerçekleştirilir. İlk adımda; Lagrange Çarpanları yöntemi kullanılarak kısıtlamalı optimizasyon problemi (2.45), eşitlik (2.46) ile gösterilen kısıtlamasız optimizasyon problemi biçiminde tekrar yazılır.

$$L_P = \frac{1}{2} \|w\|^2 - \sum_{i=1}^k \alpha_i [y_i(w' \cdot x_i + b) - 1] \quad (2.46)$$

$$= \frac{1}{2} \|w\|^2 - \sum_{i=1}^k \alpha_i y_i w' \cdot x_i - \sum_{i=1}^k \alpha_i y_i b + \sum_{i=1}^k \alpha_i$$

Burada $\alpha_i \geq 0$ değerleri Lagrange Çarpanları olarak adlandırılır. İkinci adımda ise eşitlik (2.46)'yı sadece α_i Lagrange Çarpanları'na bağlı olarak optimize edebilecek bir ikincil problem formuna dönüştürmek için Karush-Kuhn-Tucker Koşulları kullanılır (Burges, 1998; Sanei ve Chambers, 2007). Buna göre eşitlik (2.46)'nın aşağıda belirtilen şartları sağlaması gerekir:

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^k \alpha_i x_i y_i \quad (2.47)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^k \alpha_i y_i = 0 \quad (2.48)$$

Bu koşullar eşitlik (2.46)'da yerine yazılacak olursa optimizasyon probleminin ikincil problem formuna dönüşmüş biçimi:

$$\max \left(L_P = -\frac{1}{2} \sum_{i,j=1}^k \alpha_i \alpha_j y_i y_j x_i x_j + \sum_{i=1}^k \alpha_i \right) \quad (2.49)$$

olarak elde edilir. Buna bağlı sınırlamalar ise;

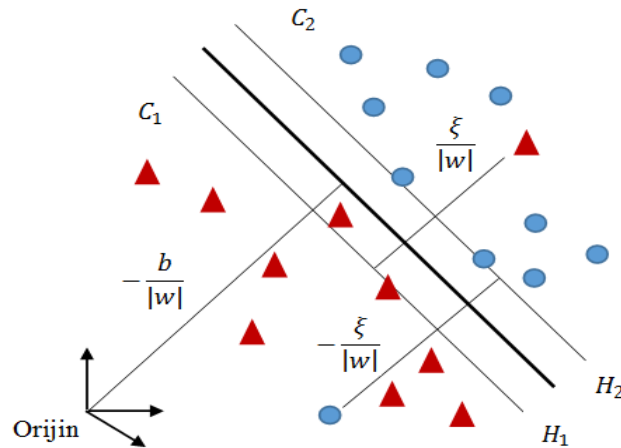
$$\sum_{i=1}^k \alpha_i y_i = 0, \quad \alpha_i \geq 0, \quad i = 1, \dots, k$$

şeklinde ifade edilir. Elde edilen bu ikincil optimizasyon problemi verilen kısıtlamalar altında karesel (ikinci dereceden) programlama yöntemleri kullanılarak çözülür. Bu problemin çözümü ile α_i değerleri bulunur. w ise eşitlik (2.47)'den elde edilir. Burada dikkat edilirse eğitim veri setinin her örneği için bir tane Lagrange çarpanı bulunmaktadır. Lagrange Çarpanı pozitif değerli olan eğitim veri seti örnekleri (x_i) destek vektörleri olarak isimlendirilirler ve H_1 ve H_2 hiper düzlemleri üzerinde yer alırlar. Lagrange Çarpanı sıfır olan diğer eğitim veri seti örnekleri ise H_1 ve H_2 hiper düzlemlerinin arka tarafında kalan örneklerdir (Burges, 1998; Sanei ve Chambers, 2007).

2.7.2.2 Lineer Olarak Belirli Bir Hata ile Ayrılabilir Durumu için DVM

Eğitim veri seti doğrusal olarak tamamen ayrılabilir durumda değilse verileri minimum hata ile sınıflandırabilecek bir çözüm algoritması kullanılır. Bu algoritma, bir önceki alt başlık altında verilmiş olan doğrusal olarak ayrılabilir verilerin sınıflandırması probleminin çözümü için kullanılan algoritmanın genelleştirilmiş bir biçimidir (Cortes ve Vapnik, 1995).

Lineer olarak tamamen ayrılamayan verilerin sınıflandırılması problemi, Şekil 2.24'te görüleceği üzere pozitif değerler alan ve sınıflandırma hatalarını ifade eden değişkenlerin (ξ_i) tanımlanması ve optimizasyon modeline eklenmesi ile çözülür (Cortes ve Vapnik, 1995).



Şekil 2.24 Doğrusal olarak belirli bir hata ile ayrılabilen veri kümesi için DVM sınıflandırıcısı

Yukarıda belirtilen ifadeler doğrultusunda doğrusal olarak tamamen ayrılabilen veriler için tanımlanan optimizasyon problemi ξ_i değişkenleri dikkate alınarak yeniden düzenlenecek olursa;

$$\min \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^k \xi_i \right) \quad (2.50)$$

olarak elde edilir. Buna bağlı sınırlamalar ise;

$$y_i(w' \cdot x_i + b) - 1 + \xi_i \geq 0, \quad \xi_i \geq 0, \quad i = 1, \dots, k$$

şeklinde ifade edilir. (2.50)'de verilen ikinci bileşenin $\sum_{i=1}^k \xi_i$ çarpanı, yanlış sınıflandırma hatasının üst sınırını belirtir. C çarpanı ise marjinin maksimum ve yanlış sınıflandırma hatasının minimuma getirilmesi arasındaki dengeyi sağlayan işlevini gören sabit bir parametredir. C parametresi dengeleyici bir işlev yapması dolayısıyla *düzenleyici parametre* olarak adlandırılır. Bu parametrenin değeri kullanıcı tarafından belirlenir. Eğer optimizasyon problemi yüksek boyutlu uzayda çözülüyorsa ikincil forma dönüştürülmelidir. Bu aşamada yapılması gereken işlemler, bir önceki alt başlık altında verilen doğrusal ayrılabilir durumdaki veriler üzerinde gerçekleştirilen işlemlerle aynıdır (Cortes ve Vapnik, 1995; Sanei ve Chambers, 2007). Bu işlemlerin burada da tekrarlanması ile optimizasyon probleminin ikincil formu;

$$\max \left(L_P = -\frac{1}{2} \sum_{i,j=1}^k \alpha_i \alpha_j y_i y_j x_i x_j + \sum_{i=1}^k \alpha_i \right) \quad (2.51)$$

olarak elde edilir (Cortes ve Vapnik, 1995; Sanei ve Chambers, 2007). Buna bağlı sınırlamalar ise;

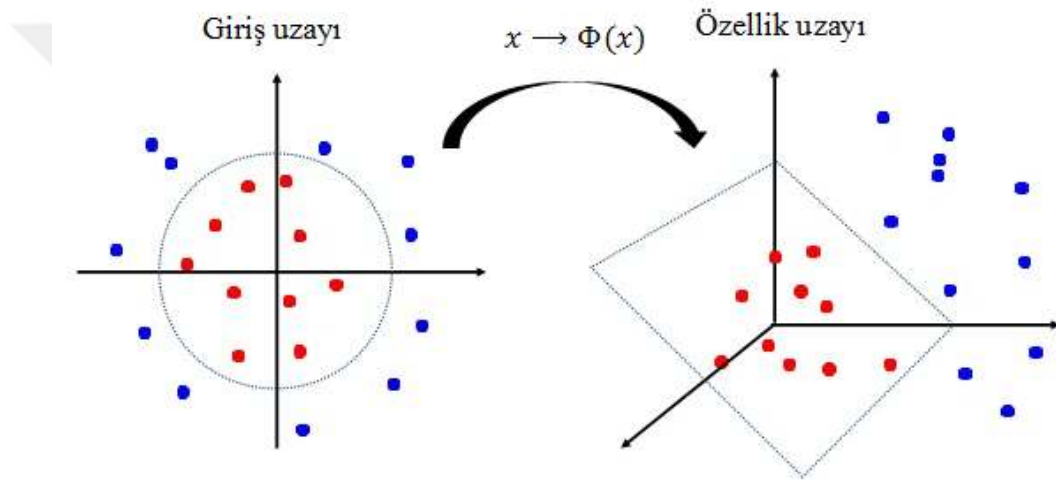
$$\sum_{i=1}^k \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, \dots, k$$

şeklinde ifade edilir. Bu problem ikinci dereceden (karesel) programlama teknikleri kullanılarak doğrusal olarak ayrılamayan veriler için en uygun DVM sınıflandırıcı parametreleri belirlenir. Burada $0 \leq \alpha_i \leq C$ aralığındaki Lagrange Çarpanları'na karşılık

gelen x_i örnekleri DVM sınıflandırıcısının destek vektörlerini temsil eder (Cortes ve Vapnik, 1995; Sanei ve Chambers, 2007).

2.7.2.3 Lineer Olmayan DVM

Doğrusal olarak tamamen veya belirli bir hata ile ayrılabilen verilerin sınıflandırılması probleminin çözümü için geliştirilen DVM algoritmalarının pratik uygulamaları oldukça kısıtlıdır. Çünkü gerçekte karşılaşılan veri kümelerinin giriş uzayındaki dağılımı Şekil 2.25'te görüleceği gibi doğrusal olmadığından bu verileri bulundukları asıl uzayda doğrusal bir ayırıcı düzlem ile ayrıştırmak pek mümkün olmamaktadır (Vapnik, 1995; Sanei ve Chambers, 2007).



Şekil 2.25 Giriş uzayının doğrusal sınıflamanın mümkün olduğu çok boyutlu uzaya dönüşümü

Bulundukları asıl uzayda doğrusal olarak ayrıştırılamayan veriler bir çekirdek fonksiyonu yardımıyla yüksek boyutlu bir özellik uzayına taşınarak o uzayda doğrusal olarak ayrıştırılabilir (Vapnik, 1995). Böylece yüksek boyutlu özellik uzayındaki çok boyutlu hiper düzlem, asıl uzayda doğrusal olmayan bir işleve karşılık gelir.

Eğitim veri setine ait d boyutlu örnekler bulundukları giriş uzayından bir Φ fonksiyonu yardımıyla daha yüksek boyutlu H özellik uzayına taşınırlar ($\Phi: \mathbb{R}^d \rightarrow H$). Bu durumda DVM'nin eğitim algoritması H uzayındaki verilerin $\Phi(x_i) \cdot \Phi(x_j)$ iç çarpımına bağlı olacaktır. Ancak yüksek boyutlu H özellik uzayındaki bu iç çarpımın hesaplanması güç olacağından bunun yerine düşük boyutlu giriş uzayında çekirdek fonksiyonları kullanılarak bu işlem gerçekleştirilir (Vapnik, 1995; Burges, 1998; Sanei ve Chambers, 2007). Kernel fonksiyonları olarak da adlandırılan ve K ile gösterilen bu çekirdek fonksiyonlarının genel ifadesi,

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad (2.52)$$

şeklindedir.

Literatürde en çok kullanılan çekirdek fonksiyonları:

Lineer: $K(x_i, x_j) = x_i \cdot x_j$

Polinomiyal: $K(x_i, x_j) = (x_i \cdot x_j + 1)^d$

Radyal: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$

EEG verileri üzerinde sınıflandırma yapılan çalışmalarda da çoğunlukla Gauss Radyal Taban Fonksiyonunun (RBF) çekirdek fonksiyon olarak kullanıldığı gözlenmektedir (Mirowski ve ark, 2009, Subaşı ve Gürsoy, 2010; Park ve ark, 2011).

2.7.3 Sınıflandırma Başarısı Ölçütleri

Sınıflandırma sonucunun başarısını değerlendirmede birtakım ölçütler kullanılmaktadır. Bu çalışmada sınıflandırma sonucunun başarısını değerlendirmede *duyarlılık*, *özgüllük*, *doğruluk* ve *Youden İndeksi* kullanıldı. Bu oranların hesaplanmasında sınıflandırma tablosundan yararlanılır (Tablo2.3).

Tablo 2.3 Sınıflandırma tablosu

		Gerçek Durum		Toplam
		Hasta	Sağlam	
Sınıflandırma Sonucu	Hasta	n_{11}	n_{12}	$n_{1.}$
	Sağlam	n_{21}	n_{22}	$n_{2.}$
	Toplam	$n_{.1}$	$n_{.2}$	n

Duyarlılık; gerçekte hasta olan bireylerin model tarafından hangi oranda saptanabildiğini gösteren bir orandır.

$$Duyarlılık = \frac{n_{11}}{n_{.1}}$$

eşitliğiyle hesaplanır (Metz, 1979).

Özgüllük; gerçekte hasta olmayan bireylerin model tarafından hangi oranda saptanabildiğini gösteren bir orandır.

$$\text{Özgüllük} = \frac{n_{22}}{n_{.2}}$$

eşitliğiyle hesaplanır (McNeil ve Adelstein, 1976).

Doğruluk; tüm bireyler içinde gerçekten negatif ve gerçekten pozitif olan birimlerin model tarafından hangi oranda saptanabildiğini gösteren bir orandır.

$$\text{Doğruluk} = \frac{n_{11} + n_{22}}{n}$$

eşitliğiyle hesaplanır (Krzanowski ve Hand, 2009).

Youden İndeksi (J); bir tanı testinin genel performansını gösteren bir istatistiktir (Schisterman ve ark, 2005).

$$J = \text{Duyarlılık} + \text{Özgüllük} - 1$$

eşitliğiyle hesaplanır (Schisterman ve ark, 2005).

Bu çalışmada yöntemlerin sınıflandırma performansları duyarlılık, özgüllük, doğruluk ve Youden İndeksi oranlarına göre karşılaştırıldı.

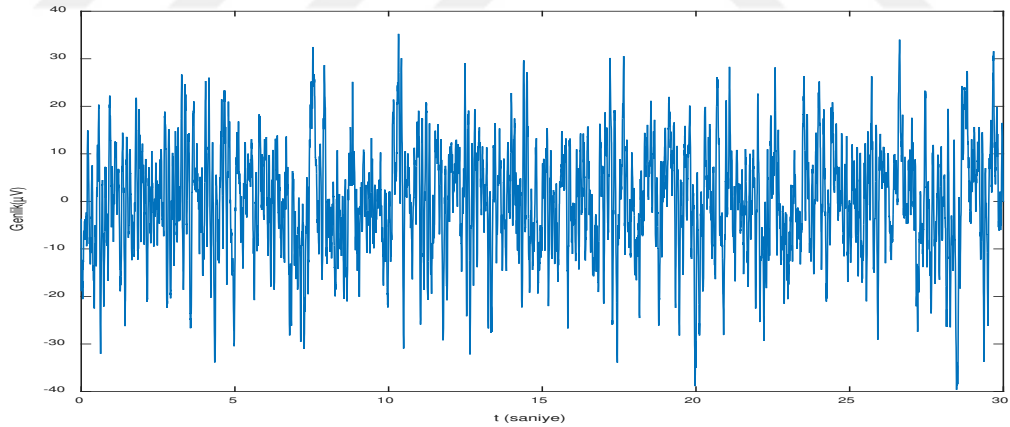
3. GEREÇ VE YÖNTEM

3.1 Verilerin Toplanması

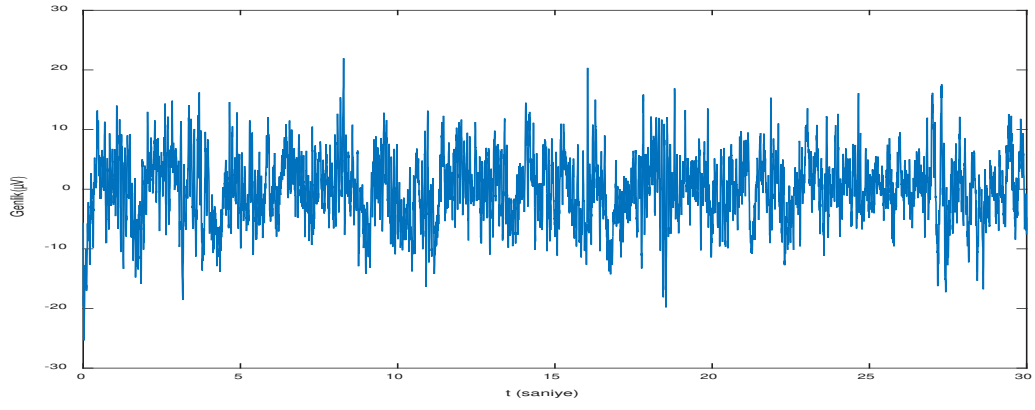
Bu çalışmada kullanılan EEG verileri, etik kurul onayı (Ek 1) alınarak, 01.04.2016 – 29.06.2016 tarihleri arasında Adnan Menderes Üniversitesi Uygulama ve Araştırma Hastanesi Nöroloji Polikliniği'nde, Amerikan EEG derneğinin klinik EEG çekimlerinde uyguladığı uluslararası 10-20 standardına uygun olarak yerleştirilen elektrotlar ile 16 kanaldan yüksek kalite ve güvenilirlikte çekim yapan ve şebeke gürültüsünü otomatik temizleyebilen Micromed EEG çekim cihazıyla kaydedilmiştir. Kayıtlar uzman hekim tarafından epilepsi tanısı konmuş 10 hasta ve 10 kontrol olmak üzere toplam 20 adet kayıttan oluşmaktadır. Her bir kayıt 30sn uzunlukta olup 256 Hz örnekleme frekansı ile kaydedilmiştir.

Kayıtlar 16 kanaldan A1 ve A2 referans noktalarına göre kaydedilmiştir. Bu kanallar; Fp1-A1, Fp2-A2, F3-A1, F4-A2, C3-A1, C4-A2, P3-A1, P4-A2, O1-A1, O2-A2, F7-A1, F8-A2, T3-A1, T4-A2, T5-A1, T6-A2'dir.

Şekil 3.1 ve Şekil 3.2'de sırasıyla epilepsi tanısı konulmuş bir hastaya ve epilepsi tanısı konulmamış bir bireye ait C3 kanalından kaydedilmiş EEG sinyalleri gösterilmektedir.



Şekil 3.1 Epilepsi tanısı konulmuş bir hastaya ait EEG sinyali



Şekil 3.2 Sağlam bir bireye ait EEG sinyali

3.2 Veri Setlerinin Oluşturulması

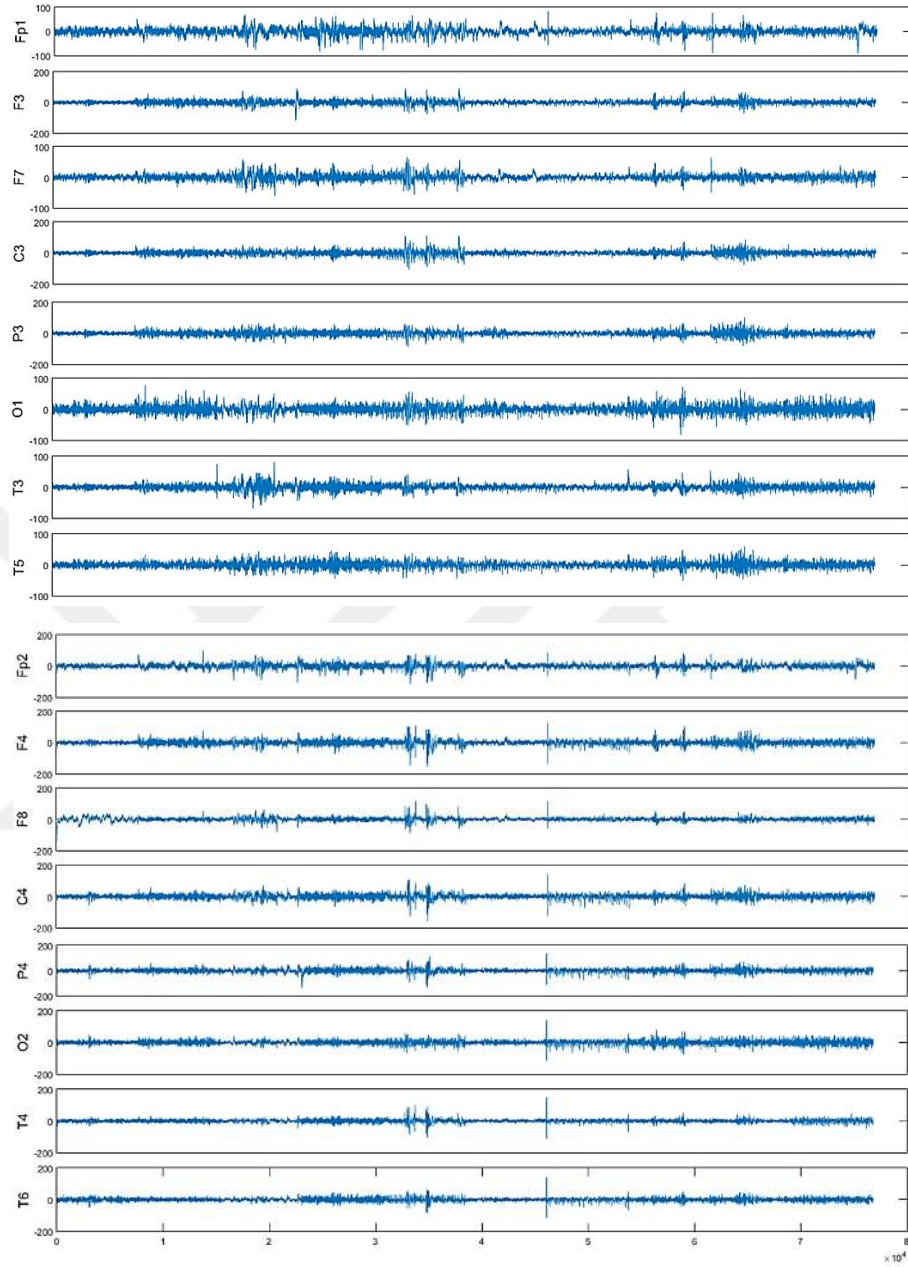
Çekim sonrası 16 kanaldan elde edilen 10 adet epileptik ve 10 adet normal EEG trasesi Şekil 3.3'teki gibi alt alta eklenerek 300'er saniye uzunlukta hasta ve sağlam bireylere ait EEG kayıtları oluşturuldu.

			Kanal 1	Kanal 2	...	Kanal 16
1. Hasta	1. s	1 2 ⋮ 256				
	⋮	⋮				
	30. s	7425 ⋮ 7680				
⋮	⋮	⋮				
10. Hasta	271. s	69121 ⋮ 69376				
	⋮	⋮				
	300. s	76545 ⋮ 76800				

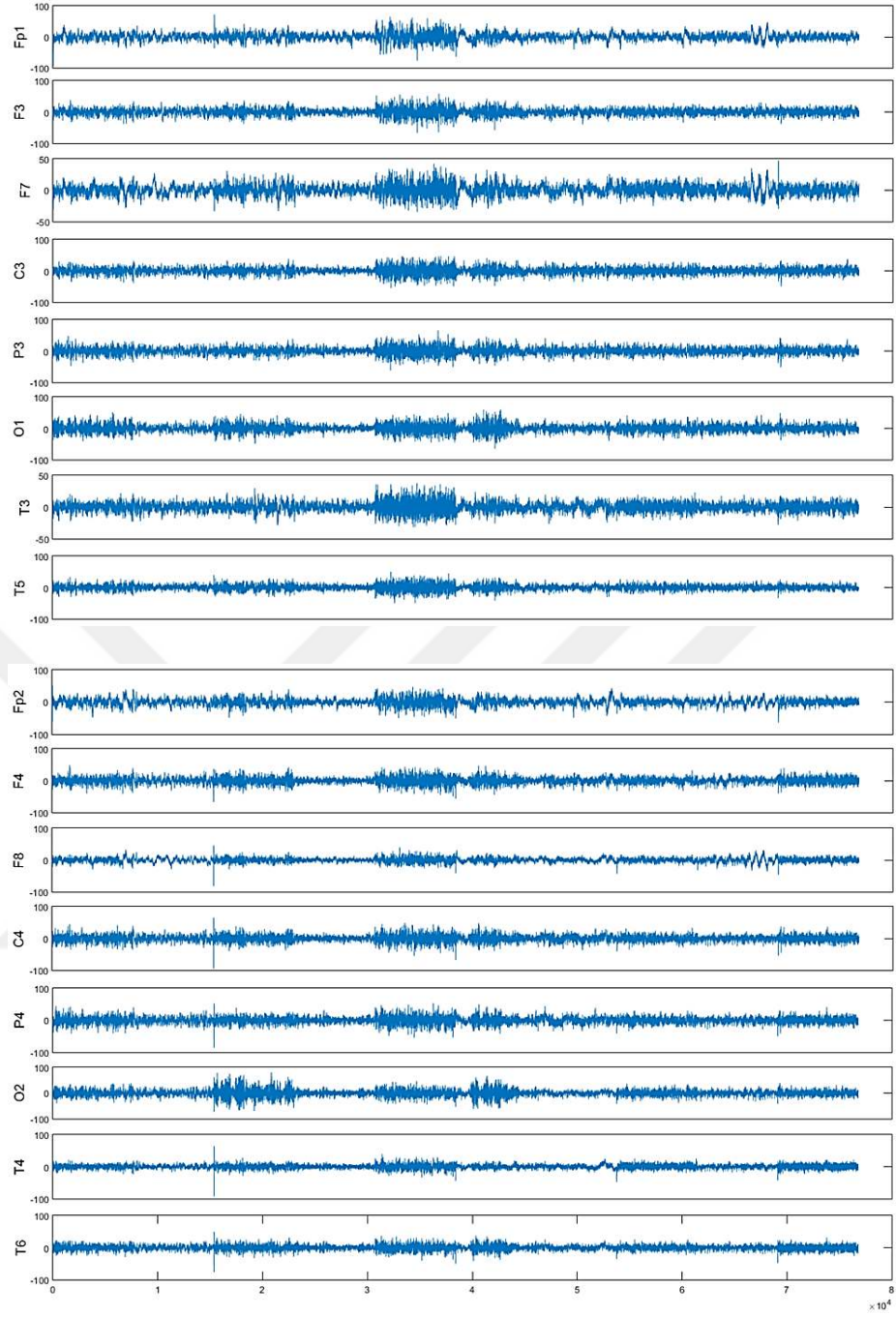
Şekil 3.3 EEG kayıtlarının birleştirilmesi

Böylelikle epileptik ve normal EEG sinyallerini temsil eden 76800x16 boyutlu iki ayrı veri matrisi elde edildi.

Hasta ve sađlam bireylere ait 300 s uzunluđundaki birleřtirilmiř EEG sinyalleri tm kanallar iin sırasıyla řekil 3.4 ve řekil 3.5'te gsterilmiřtir.



řekil 3.4 Hasta bireylere ait 300 sn uzunluđundaki birleřtirilmiř EEG sinyalleri



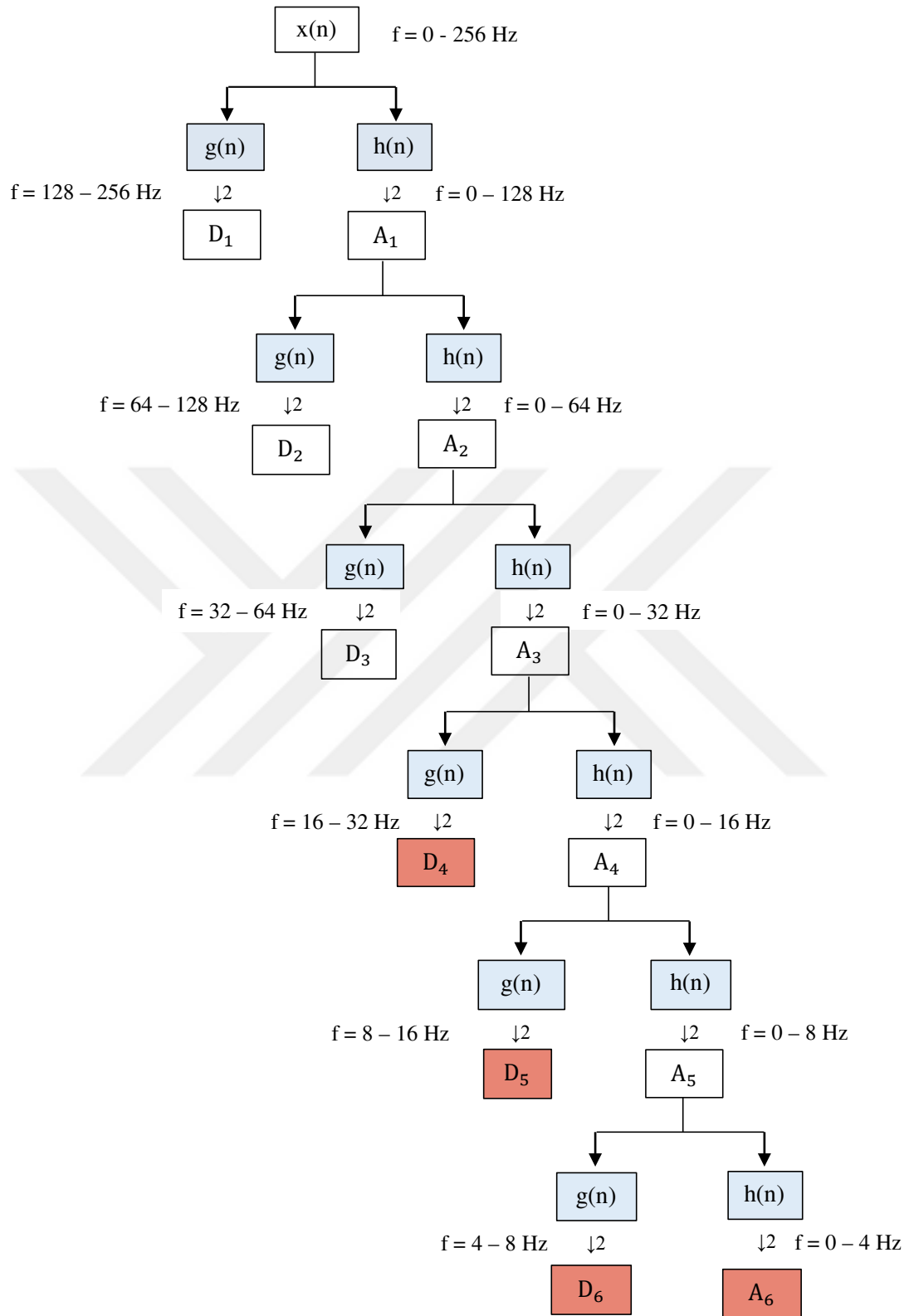
Şekil 3.5 Sağlam bireylere ait 300 sn uzunluğundaki birleştirilmiş EEG sinyalleri

3.3 Özniteliklerin Çıkarılması ve Boyutunun İndirgenmesi

Kaydedilen sinyallerin amaca yönelik kullanılabilmesi için sinyali en iyi ifade eden bazı özelliklerin belirlenerek, bu özelliklerle çalışılması gerekir. Bu özellikler literatürde öznitelik olarak adlandırılırlar. Kısaca öznitelik çıkarma işlemi, sinyallerin sınıflandırılmasını kolaylaştıracak bir forma dönüştürülmesi şeklinde ifade edilebilir. EEG verilerinden anlamlı bilgi elde edilmesi ve sınıflandırmanın başarılı bir şekilde yapılabilmesi için öznitelik vektörlerinin uygun şekilde oluşturulması önemlidir (Hyvärinen ve Oja, 2000).

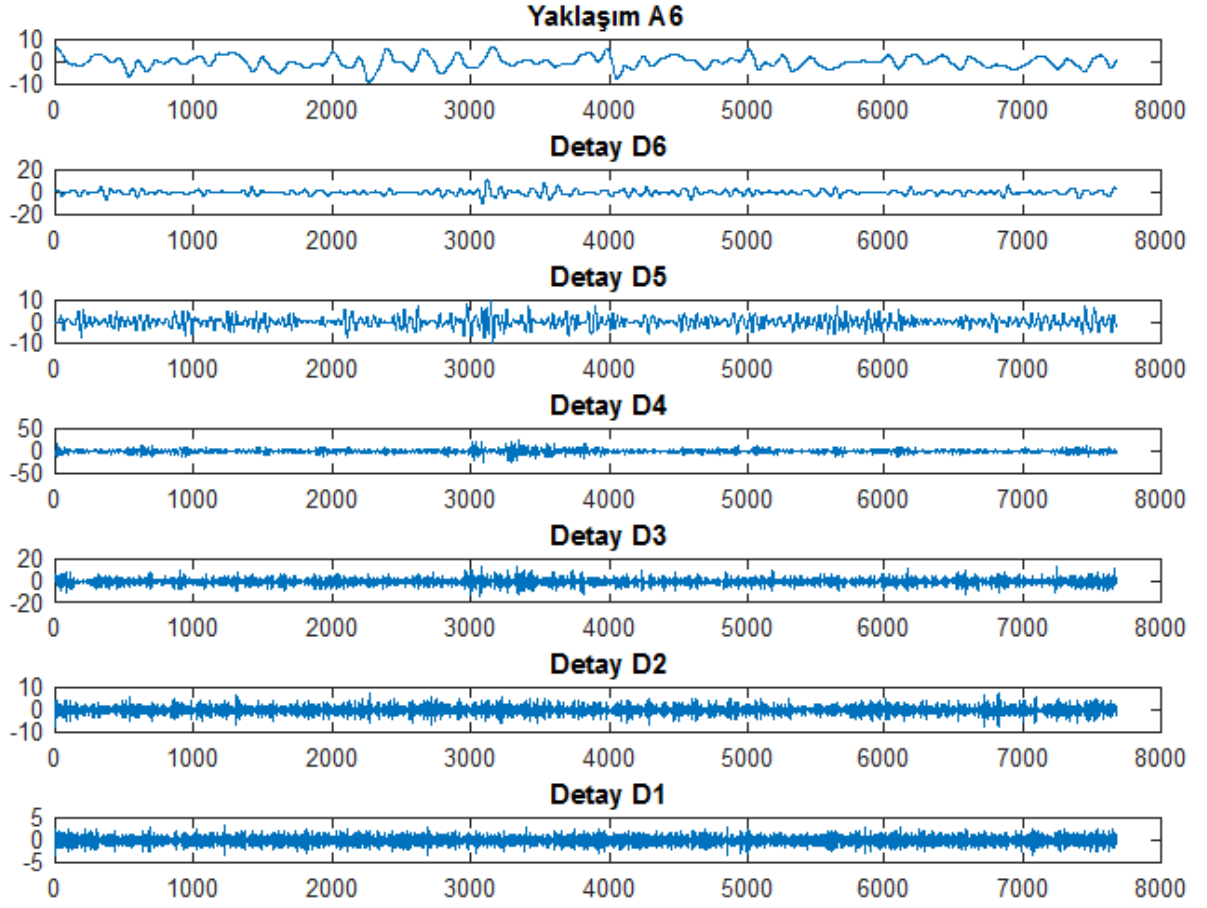
Bu çalışmada öznitelik vektörlerinin oluşturulması için EEG sinyalleri gibi durağan olmayan sinyallerin analizinde FD'ye göre daha etkin olan ve birçok çalışmada kullanılan ADD kullanıldı (Güler ve Übeyli, 2004; Übeyli, 2007, Coşkun ve İstanbullu, 2012). EEG sinyallerinin spektral analizi, ADD ile gerçekleştirildi ve sinyal hakkında önemli bilgi içeren dalgacık katsayıları sinyali temsil eden öznitelik vektörleri olarak ele alındı.

Bu çalışmada analizler farklı dalgacık fonksiyonlarıyla çok kez tekrarlandı ve elde edilen yüksek başarısından dolayı öznitelik vektörlerinin oluşturulması için Daubechies dalgacık türünün 4. dalgacığı olan Db4 (Şekil 2.17) kullanıldı. EEG sinyalleri Db4 dalgacığında 6. seviye dönüşüme tabi tutularak D1, D2, D3, D4, D5, D6 detay katsayıları ve A6 yaklaşım katsayısı elde edildi. 256 Hz'lik örnekleme frekansı göz önüne alındığında, ilk üç detay katsayısı yüksek frekans aralığına sahip olduğundan D1, D2 ve D3 katsayıları kullanılmadı. Öznitelik vektörleri oluşturulurken D4, D5, D6 detay katsayıları ve A6 yaklaşım katsayısı kullanıldı. Şekil 3.6'daki diyagramda $x(n)$ sinyaline uygulanan 6. seviye ADD sonucu elde edilen alt bantlar ve bu alt bantlara ilişkin frekans aralıkları gösterilmiştir.



Şekil 3.6 $x(n)$ sinyaline uygulanan 6. seviye ADD sonucu elde edilen alt bantlar ve bu alt bantlara ilişkin frekans aralıkları

Epileptik bir sinyalin 6. seviye ADD yöntemiyle alt bantlarına ayrıştırılması sonucunda elde edilen yaklaşım (A6) ve detay (D1-D6) alt bantları Şekil 3.7’de gösterilmektedir.



Şekil 3.7 Epileptik sinyalin 6. seviye ADD yöntemiyle alt bantlarına ayrıştırılması ile elde edilen yaklaşım ve detay alt bantları

Dalgacık dönüşümü sonucunda elde edilen detay ve yaklaşım katsayılarının frekans aralıkları Tablo 3.1’de gösterilmektedir.

Tablo 3.1 Altıncı seviye dalgacık ayrışımındaki frekans bantlarının aralıkları

Alt bantlar	Frekans Aralığı(Hz)
A6 (δ)	0 - 4
D6 (θ)	4 - 8
D5 (α)	8 - 16
D4 (β)	16 - 32
D3	32 - 64
D2	64 - 128
D1	128 - 256

Özniteliklerin elde edilmesi işleminde durağan yapıda olmayan EEG sinyallerini daha durağan bir yapıya dönüştürmek için sinyaller çerçevelere bölündü. Bu amaçla, epileptik ve normal sinyaller birer saniye uzunluğunda 300 adet çerçeveye bölündü ve ADD ile öznitelik çıkarım işlemi, her çerçeveye ayrı ayrı uygulandı. ADD ile elde edilen D4-D6, A6 dalgacık katsayıları sinyal hakkında önemli bilgi içerdiği için öznitelik vektörleri olarak ele alındı.

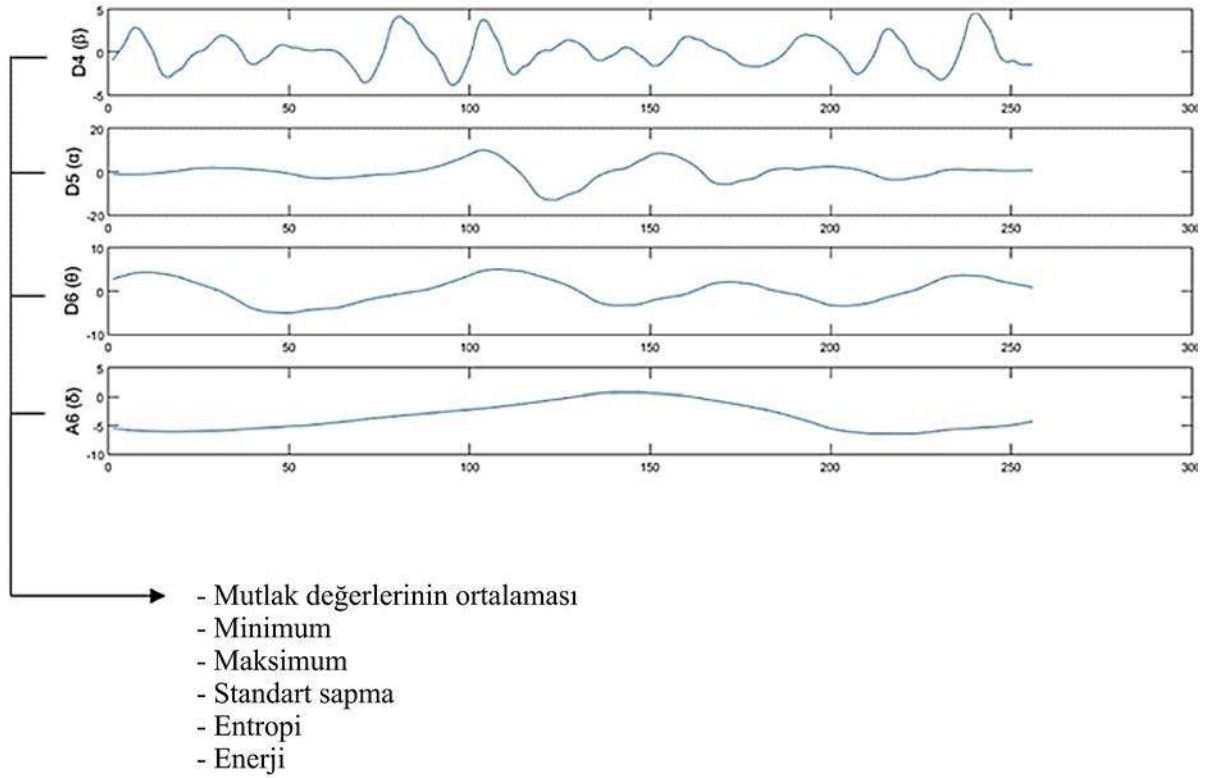
Sinyallerin özetlenmesinde en basit yöntem, analiz edilecek sinyallerin karakteristiklerine uygun istatistiksel işlemlerin uygulanarak sinyalden birkaç istatistik elde edilmesi ve bu istatistiklerin incelenmesidir. Öznitelik vektörlerini oluşturmak için literatürde önerilmiş birçok istatistik olsa da EEG sinyalleri için en sık kullanılan istatistikler ortalama, maksimum, minimum, standart sapma, entropi ve enerji değerleridir (Kandaswamy ve ark, 2004; Li ve Wen, 2009; Subaşı ve Gürsoy, 2010; Ersöz ve Özşen, 2011; Ahammad ve ark, 2014). Bazen elde edilen öznitelikler veriyi o kadar iyi temsil eder ki kompleks bir sınıflandırma yöntemi yerine basit bir eşik değeri belirlemek bile sınıfları ayırt edebilmek için yeterlidir. Bu şekilde basit bir sınıflandırma yöntemi kullanarak yakalanan yüksek başarı, genellikle güçlü bir özellik çıkartma aşamasının ürünü olarak yorumlanır (Orhan, 2011).

Öznitelik vektörlerinin boyutlarının azaltılması ve sınıflandırma performansının artırılması için dalgacık katsayılarından aşağıdaki istatistikler hesaplanarak öznitelik vektörleri elde edildi.

1. Her bir alt banttaki katsayıların mutlak değerlerinin ortalaması ($\overline{|x|} = \frac{\sum |x_i|}{n}$)
2. Her bir alt banttaki katsayıların minimumu ($\min(x_i)$)
3. Her bir alt banttaki katsayıların maksimumu ($\max(x_i)$)

4. Her bir alt banttaki katsayıların standart sapması $\left(s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}\right)$
5. Her bir alt banttaki katsayıların entropisi ($Entropi = -\sum p_x \log p_x$)
6. Her bir alt banttaki katsayıların enerjisi ($Enerji = \sum |x_i|^2$)

Şekil 3.8’de gösterildiği gibi her çerçeve için D4, D5, D6 ve A6 dalgacık katsayıları elde edilerek, bu katsayılar için istatistiksel özellikler çıkarıldı.



Şekil 3.8 Bir saniyelik çerçeveden elde edilen D4, D5, D6 ve A6 alt bantları ve bu alt bantlardan hesaplanan istatistikler

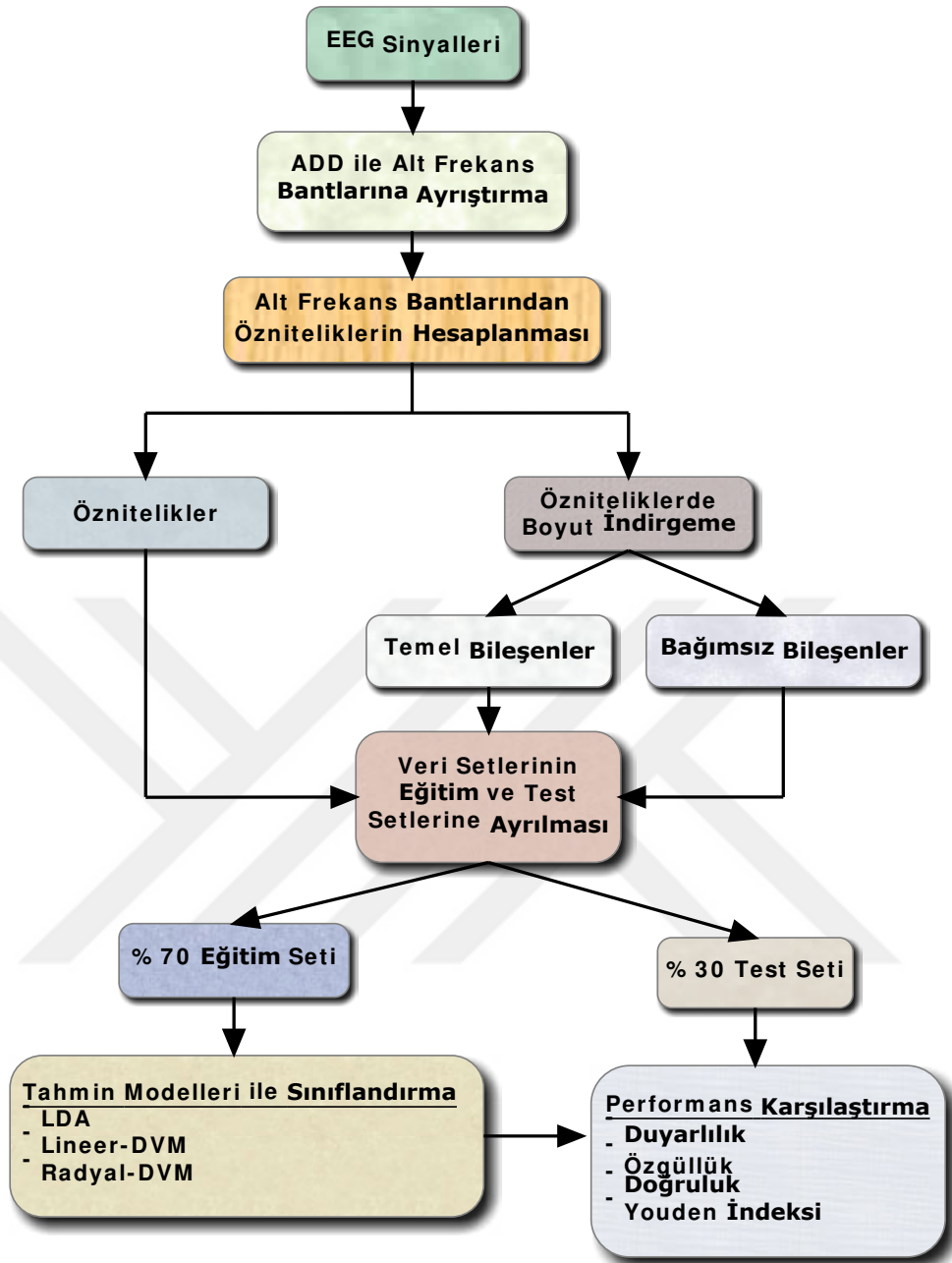
Bu durumda her bir kanaldan hem normal hem de epileptik sinyal için ((3 Detay alt bandı + 1 Yaklaşım alt bandı) x (6 İstatistiksel Özellik)) 24 adet öz nitelik elde edildi. Böylece hasta ve sağlam bireylere ilişkin her bir kanaldan 300x24 boyutlu öz nitelik matrisleri elde edildi. Örnek olarak, C3 kanalından elde edilen öz nitelik matrisi Şekil 3.9’daki gibi oluşturuldu.

		C3 kanalı							
		A6					D6	D5	D4
		Ort.	Min.	Max.	Std.	Enerji	Entropi	...	...
1.denek	1								
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	30								
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10.denek	271								
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	300								

Şekil 3.9 Öznitelik matrisi

Uygulamalarda ham veriyi temsil edecek özniteliklerin çıkarılması, sınıflandırıcı başarısını artırırken, boyutun daha da düşürülmesiyle sınıflandırma işlem yükünü de azaltmış olur (Uyar ve ark, 2008). Bu çalışmada, her bir kanal ADD ile alt frekans bantlarına ayrıştırıldı ve bu frekans bantlarından öznitelikler elde edildi. Elde edilen özniteliklerin boyutunun azaltılması için TBA ve BBA yöntemleri kullanıldı. Veri setlerinin oluşturulması, özniteliklerin çıkarılması ve boyut indirgeme MATLAB R2015a deneme sürümü, sınıflandırma analizleri ise R-3.3.1 for Windows programı ile yapıldı.

Sınıflandırmalarda kullanılan tüm öznitelik matrisleri için eğitim ve test veri setleri aynı noktalara karşılık gelecek şekilde %70'i eğitim (n=420) ve %30'u test (n=180) verisi olmak üzere iki parçaya rastgele olarak ayrıldı. Uygulamada yapılan işlem süreci Şekil 3.10'da verildi.



Şekil 3.10 İşlem sürecinin akış diyagramı

4. BULGULAR

4.1 Özniteliklerin TBA ve BBA ile Boyutlarının İndirgenmesi

TBA ile boyut indirgemenin, sınıflandırma performansına etkisini belirlemek amacıyla 16 kanaldan elde edilen 384 adet öznitelik öz değeri 1'den büyük olan ve toplam değişkenliğin %92,475'ini açıklayan 34 adet temel bileşene indirgendir. Böylece temel bileşenlerden oluşan 600x34 boyutlu yeni öznitelik matrisi elde edildi. Tablo 4.1'de TBA ile elde edilen temel bileşenlerin öz değerleri, varyans ve kümülatif varyans yüzdeleri verildi.

Tablo 4.1 Temel bileşenlerin öz değerleri, varyans ve kümülatif varyans yüzdeleri

Temel Bileşenler	Öz değer	Varyans (%)	Kümülatif varyans (%)
1	188,882	49,188	49,188
2	39,557	10,301	59,489
3	15,752	4,102	63,591
4	11,864	3,090	66,681
5	11,250	2,930	69,610
6	10,275	2,676	72,286
7	8,094	2,108	74,394
8	5,982	1,558	75,952
9	5,918	1,541	77,493
10	5,247	1,366	78,859
11	4,805	1,251	80,110
12	4,078	1,062	81,172
13	3,600	0,937	82,110
14	3,326	0,866	82,976
15	3,191	0,831	83,807
16	2,784	0,725	84,532
17	2,767	0,720	85,253
18	2,605	0,678	85,931
19	2,305	0,600	86,531
20	2,246	0,585	87,116
21	2,075	0,540	87,657
22	2,035	0,530	88,187
23	1,880	0,490	88,676
24	1,755	0,457	89,133
25	1,625	0,423	89,556
26	1,587	0,413	89,970
27	1,452	0,378	90,348
28	1,334	0,347	90,695
29	1,294	0,337	91,032
30	1,265	0,330	91,362
31	1,132	0,295	91,657
32	1,070	0,279	91,935
33	1,058	0,275	92,211
34	1,015	0,264	92,475

BBA ile boyut indirgemenin, sınıflandırma performansına etkisini belirlemek amacıyla 384 adet öznitelik, öz değeri 1'den büyük olan ve toplam değişkenliğin %92,662'sini açıklayan 30 adet bağımsız bileşene indirgendi. Böylece bağımsız bileşenlerden oluşan 600x30 boyutlu yeni öznitelik matrisi elde edildi. Çalışmamızda BBA, Gauss fonksiyonunun kullanıldığı FastICA algoritması ile yapıldı. Tablo 4.2'de BBA ile elde edilen bağımsız bileşenlerin öz değerleri, varyans ve kümülatif varyans yüzdeleri verildi.

Tablo 4.2 Bağımsız bileşenlerin öz değerleri, varyans ve kümülatif varyans yüzdeleri

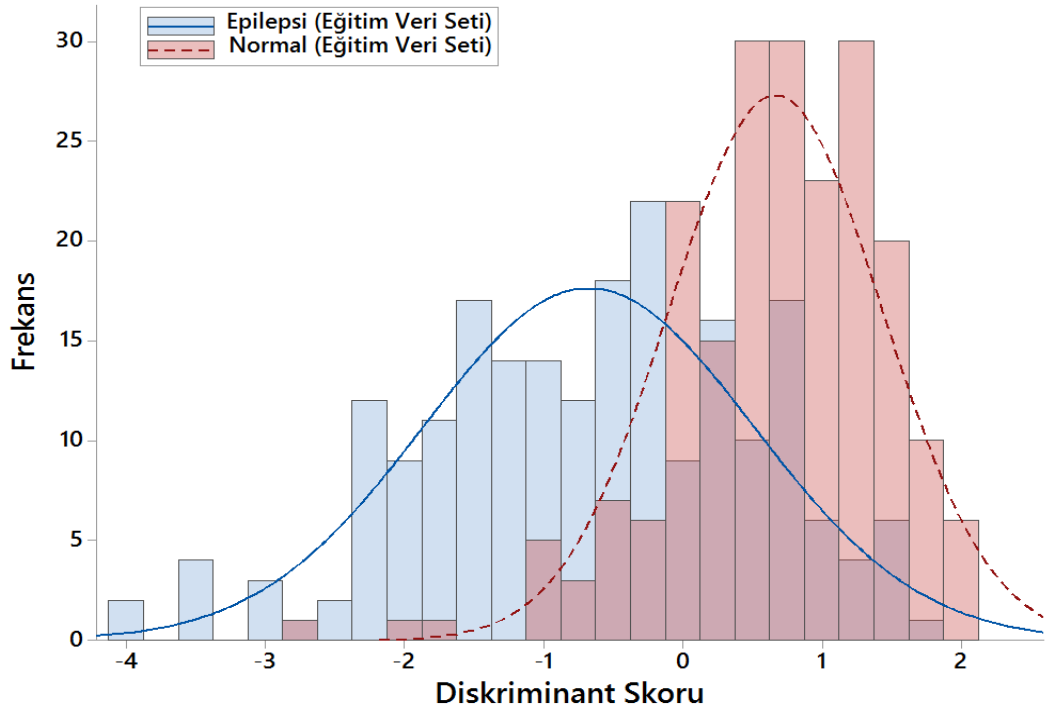
Bağımsız Bileşenler	Öz değer	Varyans (%)	Kümülatif varyans (%)
1	185,224	52,347	52,347
2	27,538	7,782	60,129
3	14,171	4,005	64,134
4	13,303	3,759	67,893
5	11,189	3,162	71,055
6	9,400	2,656	73,711
7	6,726	1,901	75,612
8	6,122	1,730	77,342
9	5,669	1,602	78,944
10	4,939	1,396	80,340
11	4,568	1,291	81,631
12	3,739	1,056	82,687
13	3,403	0,961	83,648
14	3,121	0,882	84,530
15	2,967	0,838	85,368
16	2,641	0,746	86,114
17	2,595	0,733	86,887
18	2,454	0,693	87,580
19	2,036	0,575	88,155
20	1,982	0,560	88,715
21	1,896	0,535	89,250
22	1,753	0,495	89,745
23	1,582	0,447	90,192
24	1,512	0,427	90,619
25	1,410	0,398	91,017
26	1,295	0,366	91,383
27	1,250	0,353	91,736
28	1,168	0,330	92,066
29	1,089	0,307	92,373
30	1,025	0,289	92,662

BBA sonucunda sınıflandırılmada kullanılan 600x30 boyutlu yeni öznitelik matrisi elde edildi. Böylelikle 600x384 boyutlu öz nitelik matrisi, TBA ile elde edilen 600x34 ve BBA ile elde edilen 600x30 boyutlu yeni öz nitelik matrisleri olmak üzere toplam 3 farklı öz nitelik matrisinin her biri LDA, lineer ve radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM yöntemleri ile ayrı ayrı sınıflandırıldı.

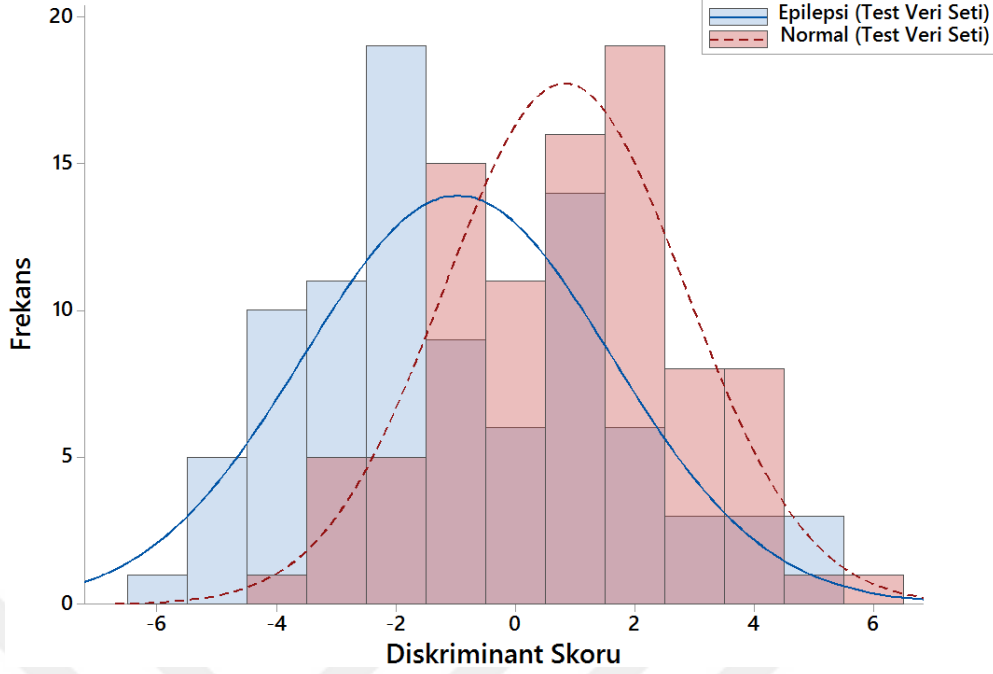
4.2 LDA Sınıflandırma Sonuçları

Özniteliklerden, temel bileşenlerden ve bağımsız bileşenlerden hesaplanan LDA sonuçları (diskriminant fonksiyonu, diskriminant skorları, kanonik korelasyon, özdeğer, Wilks' Lambda, duyarlılık, özgüllük, doğruluk ve Youden İndeksi) karşılaştırıldı.

LDA ile özniteliklerden elde edilen diskriminant skorlarının histogram grafiği eğitim ve test veri setleri için sırasıyla Şekil 4.1 ve Şekil 4.2'de verildi.



Şekil 4.1 Özniteliklerin diskriminant skorları histogram grafiği



Şekil 4.2 Özniteliklerin diskriminant skorları histogram grafiği

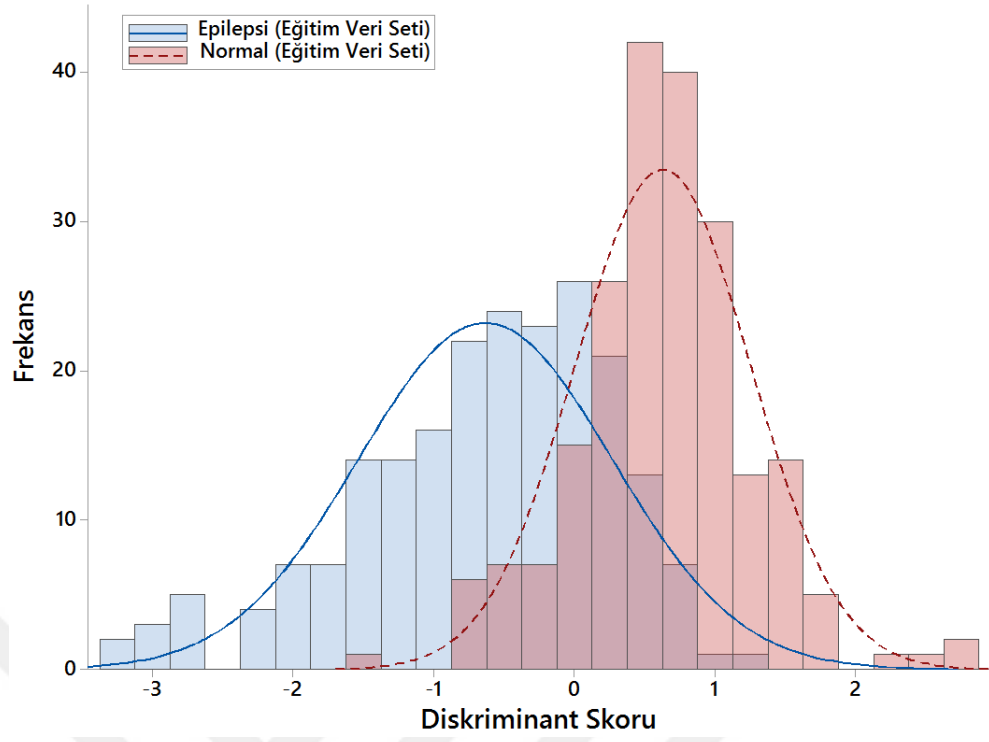
Özniteliklerden elde edilen diskriminant fonksiyonunun önemliliğine ilişkin kanonik korelasyon 0,559, öz değer 0,454 ve Wilks' Lamda istatistiği 0,688 olarak bulundu. Kanonik korelasyon değerine göre model istatistiksel olarak anlamlı bulundu ($p < 0,001$) ve kanonik korelasyona göre model, bağımlı değişkendeki varyansın %31,2'sini açıklamaktaydı.

Eğitim ve test veri setlerinin LDA ile yapılan sınıflandırma sonuçlarına ilişkin duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.3'te verildi.

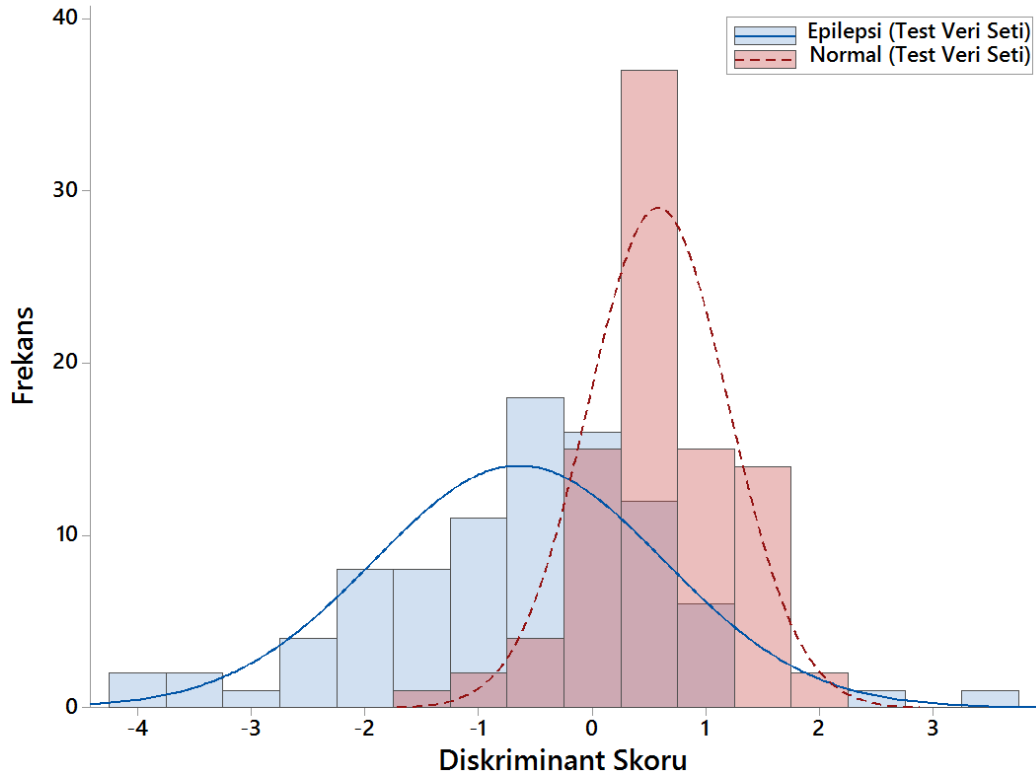
Tablo 4.3 Özniteliklere ait eğitim ve test veri setlerinin LDA sınıflandırma sonuçları

	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	69,0	83,3	76,2	52,3
Test Veri Seti	57,8	63,3	60,6	21,1

Temel bileşenlerden LDA ile elde edilen diskriminant skorlarının histogram grafiği eğitim ve test veri setleri için sırasıyla Şekil 4.3 ve Şekil 4.4'te verildi.



Şekil 4.3 Temel bileşenlerin diskriminant skorları histogram grafiği



Şekil 4.4 Temel bileşenlerin diskriminant skorları histogram grafiği

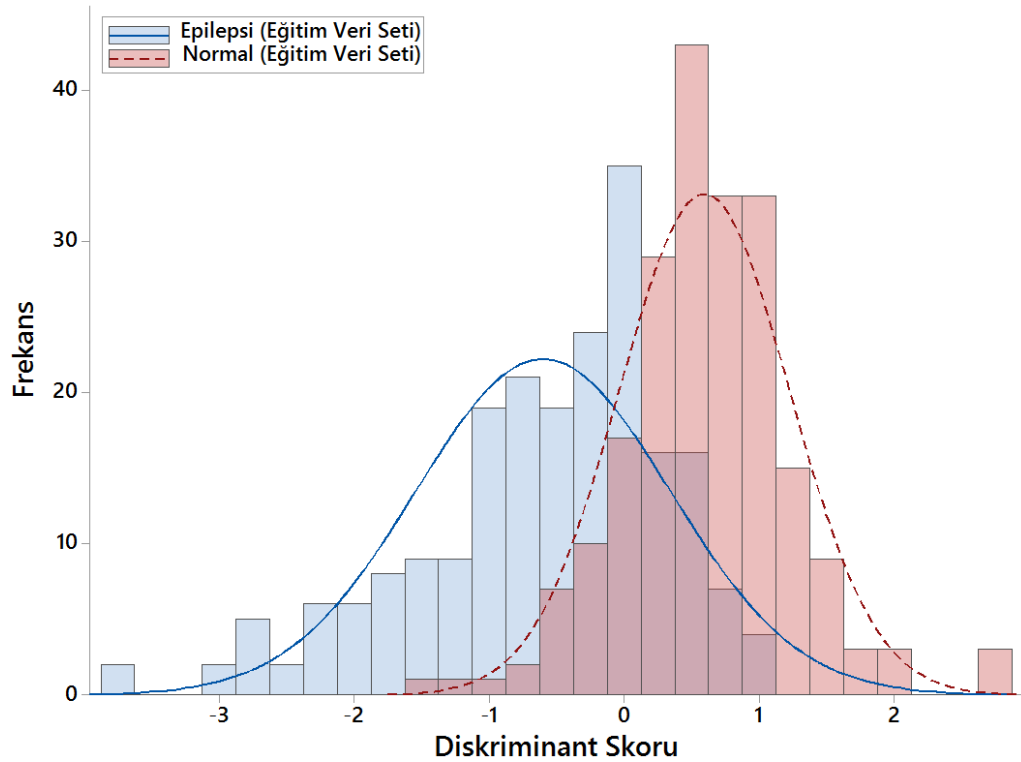
Temel bileşenlerden elde edilen diskriminant fonksiyonunun önemliliğine ilişkin kanonik korelasyon 0,632, öz değer 0,664 ve Wilks' Lamda istatistiği 0,601 olarak bulundu. Kanonik korelasyon değerine göre model istatistiksel olarak anlamlı bulundu ($p < 0,001$) ve kanonik korelasyona göre model, bağımlı değişkendeki varyansın %39,9'unu açıklamaktaydı.

Temel bileşenler için eğitim ve test veri setlerinin LDA ile yapılan sınıflandırma sonuçlarına ilişkin duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.4'te verildi.

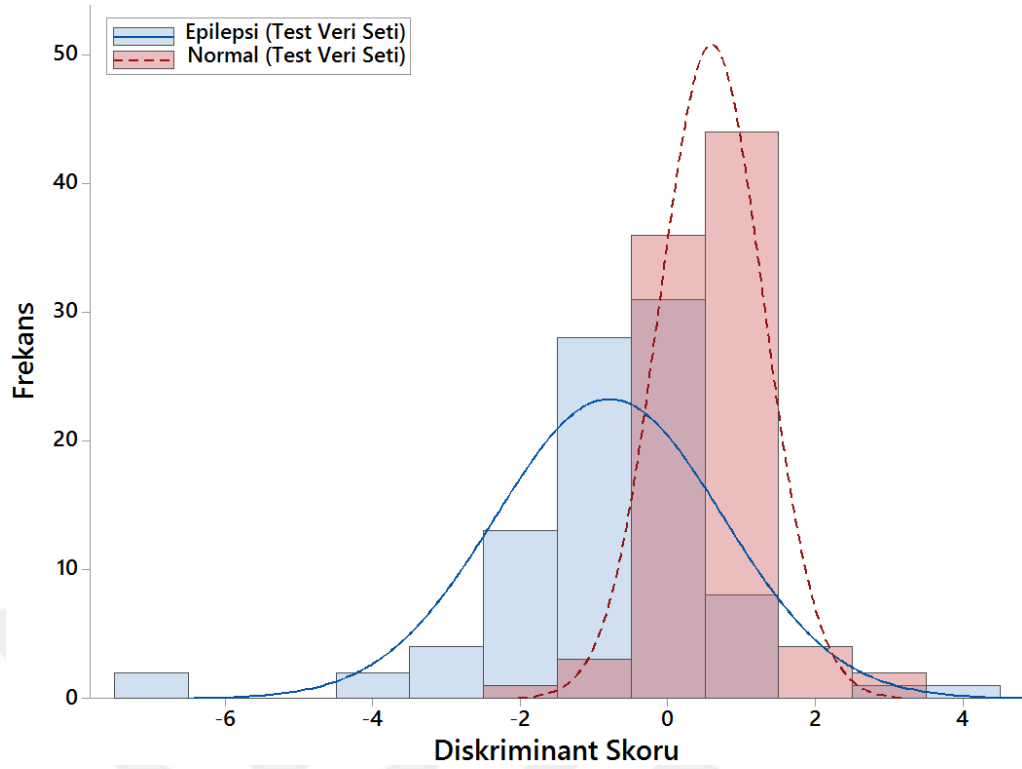
Tablo 4.4 Temel bileşenlere ait eğitim ve test veri setlerinin LDA sınıflandırma sonuçları

	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	73,8	85,7	79,8	59,5
Test Veri Seti	67,8	90,0	78,9	57,8

Bağımsız bileşenlerden LDA ile elde edilen diskriminant skorlarının histogram grafiği eğitim ve test veri setleri için sırasıyla Şekil 4.5 ve Şekil 4.6'da verildi.



Şekil 4.5 Bağımsız bileşenlerin diskriminant skorları histogram grafiği



Şekil 4.6 Bağımsız bileşenlerin diskriminant skorları histogram grafiği

Bağımsız bileşenlerden elde edilen diskriminant fonksiyonunun önemliliğine ilişkin kanonik korelasyon 0,598, öz değer 0,557 ve Wilks' Lamda istatistiği 0,642 olarak bulundu. Kanonik korelasyon değerine göre model istatistiksel olarak anlamlı bulundu ($p < 0,001$) ve kanonik korelasyona göre model, bağımlı değişkendeki varyansın %35,8'ini açıklamaktaydı.

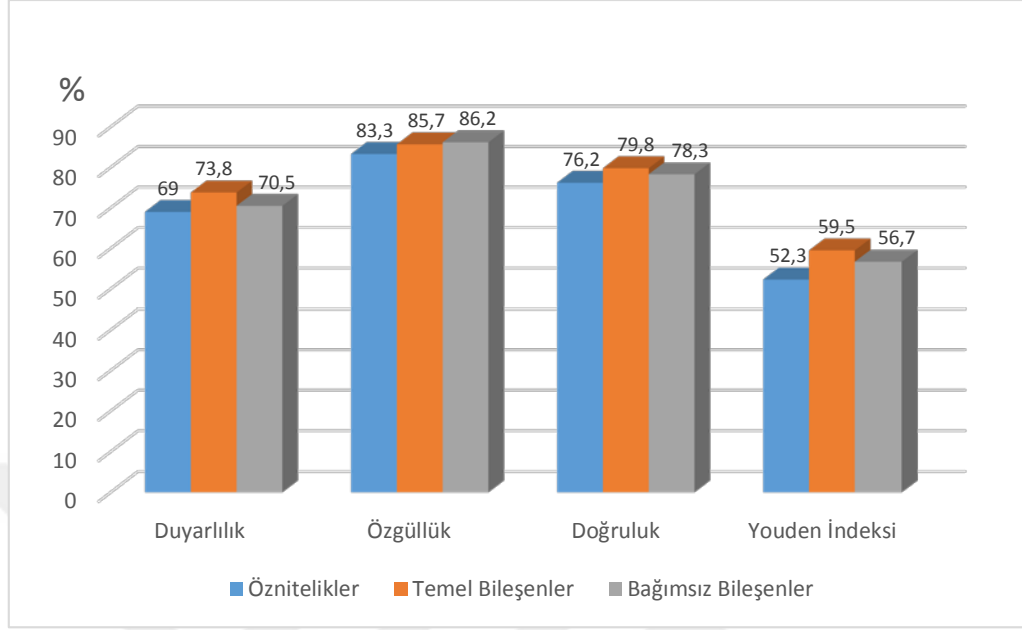
Bağımsız bileşenler için eğitim ve test veri setlerinin LDA ile yapılan sınıflandırma sonuçlarına ilişkin duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.5'te verildi.

Tablo 4.5 Bağımsız bileşenlere ait eğitim ve test veri setlerinin LDA sınıflandırma sonuçları

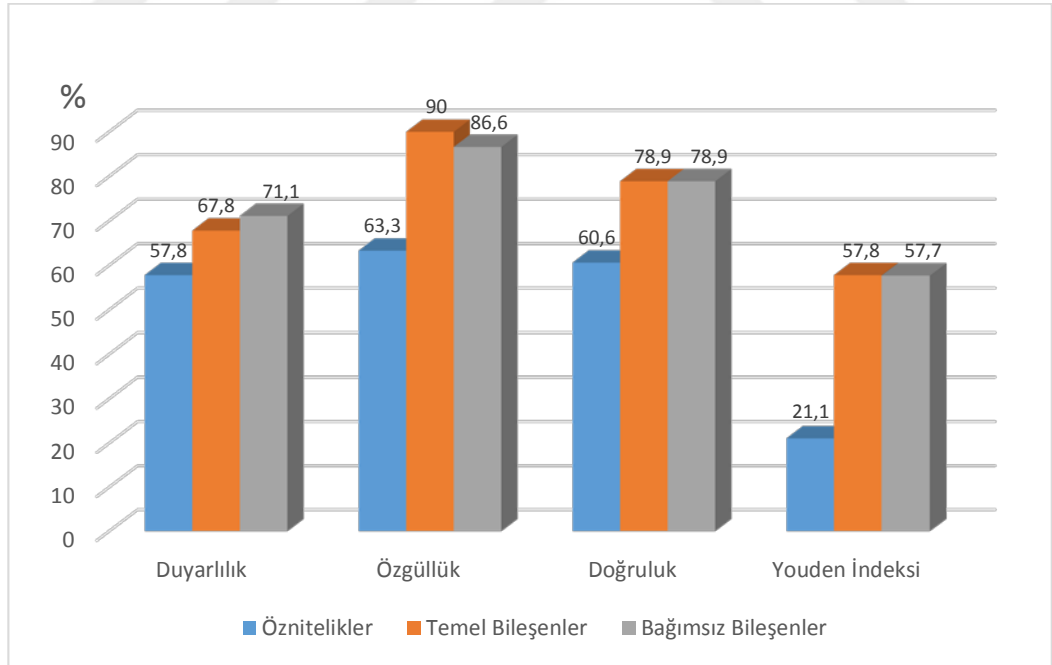
	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	70,5	86,2	78,3	56,7
Test Veri Seti	71,1	86,6	78,9	57,7

LDA ile yapılan sınıflandırma sonuçlarına göre eğitim veri seti için en yüksek duyarlılık (%73,8), doğruluk (%79,8) ve Youden İndeksi (%59,5) temel bileşenlerden elde edilirken, en yüksek özgüllük (%86,2) bağımsız bileşenlerden elde edildi (Şekil 4.7). Test veri setinde ise en yüksek duyarlılık (%71,1) bağımsız bileşenlerden, en yüksek özgüllük (%90,0) ve Youden

İndeksi (%57,8) temel bileşenlerden elde edilirken en yüksek en yüksek doğruluk oranı (%78,9) hem temel bileşenler hem de bağımsız bileşenlerden elde edildi (Şekil 4.8).



Şekil 4.7 Eğitim veri setlerinin LDA ile yapılan sınıflandırma sonuçlarının karşılaştırılması



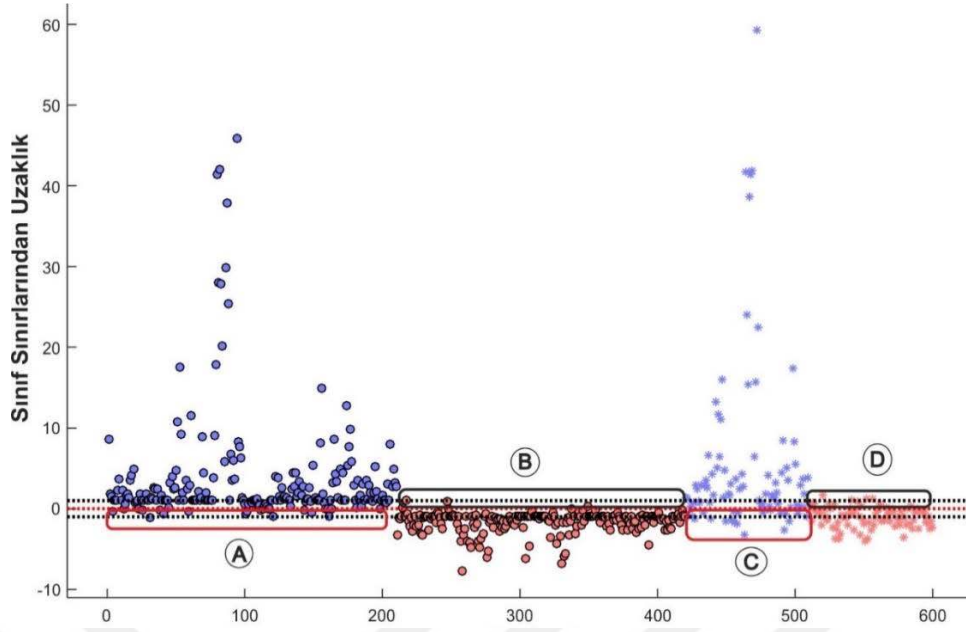
Şekil 4.8 Test veri setlerinin LDA ile yapılan sınıflandırma sonuçlarının karşılaştırılması

4.3 DVM Sınıflandırma Sonuçları

Bu bölümde özniteliklerden, temel bileşenlerden ve bağımsız bileşenlerden hesaplanan DVM sonuçları karşılaştırıldı. DVM ile yapılan sınıflandırmada hem lineer (Lineer-DVM) hem de radyal (Radyal-DVM) tabanlı çekirdek fonksiyonları kullanıldı. Bu fonksiyonlar için gerekli olan genişleme parametresi γ ve düzenleme parametresi C için en iyi değerler iki parametre için farklı değerlerin aralık tarama (grid-search) yöntemiyle denenmesi sonucunda elde edildi. Parametreler 10-kat çapraz geçerlilik yöntemi ile kontrol edildi ve en iyi çapraz doğrulama ayarını veren parametreler, en iyi parametreler olarak seçildi.

4.3.1 Lineer-DVM Sınıflandırma Sonuçları

Özniteliklerle yapılan sınıflandırmada eğitim ve test veri setleri için özniteliklerin sınıf sınırlarından uzaklıkları Şekil 4.9'da gösterildi. Şekil 4.9'da A bölgesine düşen noktalar eğitim veri setinde yanlış sınıflandırılan epileptik sinyalleri ve B bölgesine düşen noktalar ise eğitim veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Benzer şekilde C bölgesine düşen noktalar test veri setinde yanlış sınıflandırılan epileptik sinyalleri ve D bölgesine düşen noktalar ise test veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Şekil 4.9'da test veri setindeki epileptik ve normal sinyalleri temsil eden noktaların yanlış sınıflandırma oranlarının eğitim veri setindeki yanlış sınıflandırma oranlarına göre daha yüksek olduğu görülmektedir.



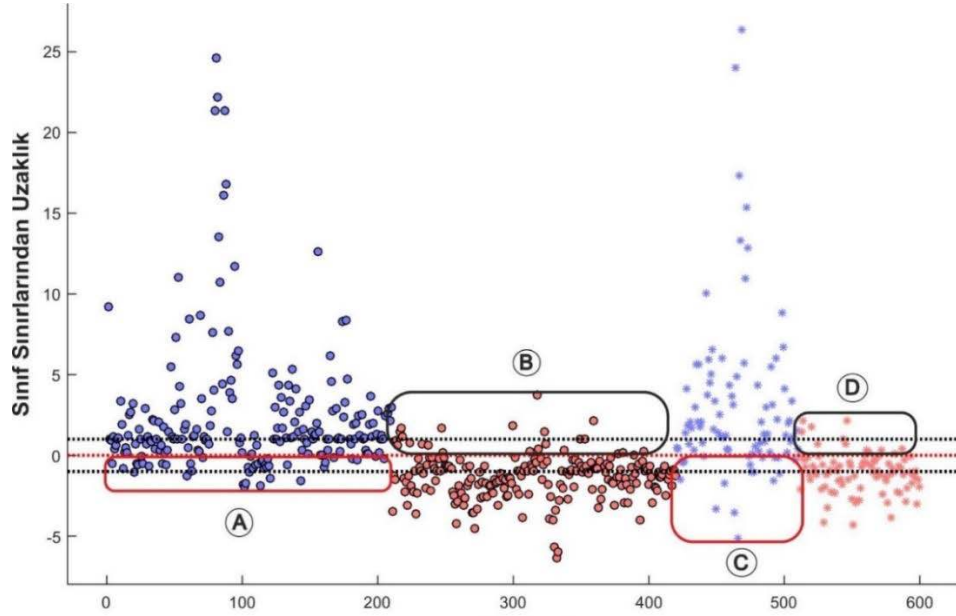
Şekil 4.9 Özniteliklerle yapılan sınıflandırmada eğitim ve test veri setleri için özniteliklerin sınıf sınırlarından uzaklıkları

Eğitim ve test veri setlerinin Lineer-DVM ile yapılan sınıflandırma sonuçlarına ilişkin düzenleme parametresi, duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.6’da verildi.

Tablo 4.6 Özniteliklere ait eğitim ve test veri setlerinin Lineer-DVM sınıflandırma sonuçları

	C	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	0,1	91,4	98,1	94,7	89,5
Test Veri Seti	-	77,7	81,1	79,4	58,8

Temel bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için temel bileşenlerin sınıf sınırlarından uzaklıkları Şekil 4.10’da gösterildi. Şekil 4.10’da A bölgesine düşen noktalar eğitim veri setinde yanlış sınıflandırılan epileptik sinyalleri ve B bölgesine düşen noktalar ise eğitim veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Benzer şekilde C bölgesine düşen noktalar test veri setinde yanlış sınıflandırılan epileptik sinyalleri ve D bölgesine düşen noktalar ise test veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Şekil 4.10’da test veri setindeki epileptik ve normal sinyalleri temsil eden noktaların yanlış sınıflandırma oranlarının eğitim veri setindeki yanlış sınıflandırma oranlarına benzer olduğu görülmektedir.



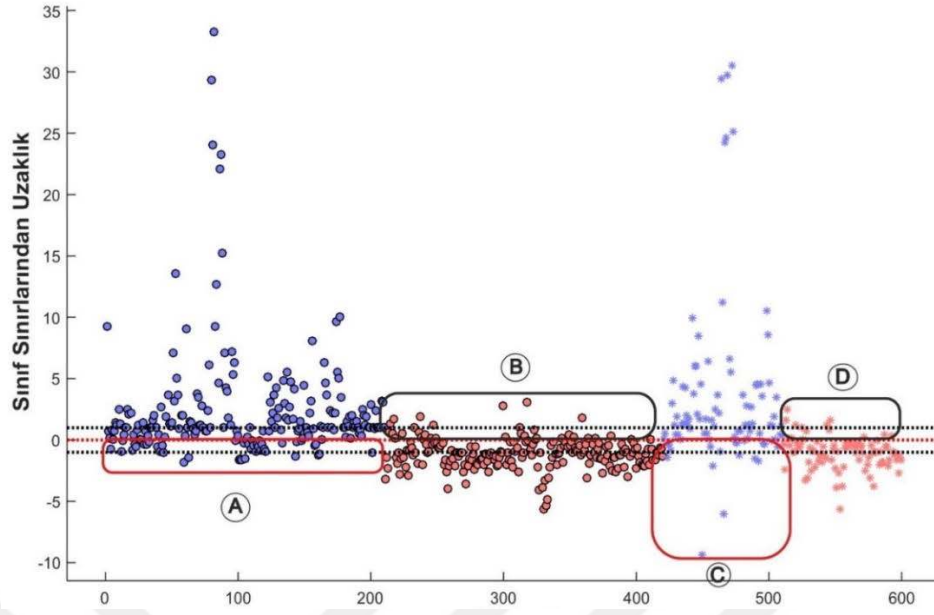
Şekil 4.10 Temel bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için temel bileşenlerin sınıf sınırlarından uzaklıkları

Eğitim ve test veri setlerinin Lineer-DVM ile yapılan sınıflandırma sonuçlarına ilişkin düzenleme parametresi, duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.7’de verildi.

Tablo 4.7 Temel bileşenlere ait eğitim ve test veri setlerinin Lineer-DVM sınıflandırma sonuçları

	C	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	1	80,0	88,6	84,3	68,6
Test Veri Seti	-	78,9	85,6	82,2	64,5

Bağımsız bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için bağımsız bileşenlerin sınıf sınırlarından uzaklıkları Şekil 4.11’de gösterildi. Şekil 4.11’de A bölgesine düşen noktalar eğitim veri setinde yanlış sınıflandırılan epileptik sinyalleri ve B bölgesine düşen noktalar ise eğitim veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Benzer şekilde C bölgesine düşen noktalar test veri setinde yanlış sınıflandırılan epileptik sinyalleri ve D bölgesine düşen noktalar ise test veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Şekil 4.11’de test veri setindeki epileptik ve normal sinyalleri temsil eden noktaların yanlış sınıflandırma oranlarının eğitim veri setindeki yanlış sınıflandırma oranlarına benzer olduğu görülmektedir.



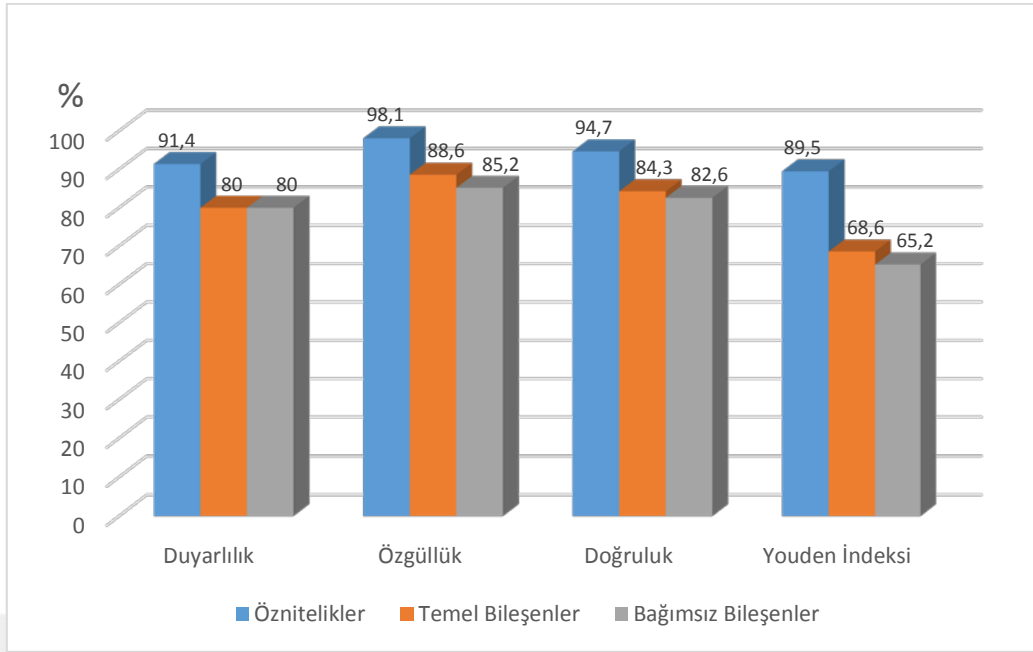
Şekil 4.11 Bağımsız bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için bağımsız bileşenlerin sınıf sınırlarından uzaklıkları

Eğitim ve test veri setlerinin Lineer-DVM ile yapılan sınıflandırma sonuçlarına ilişkin düzenleme parametresi, duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.8’de verildi.

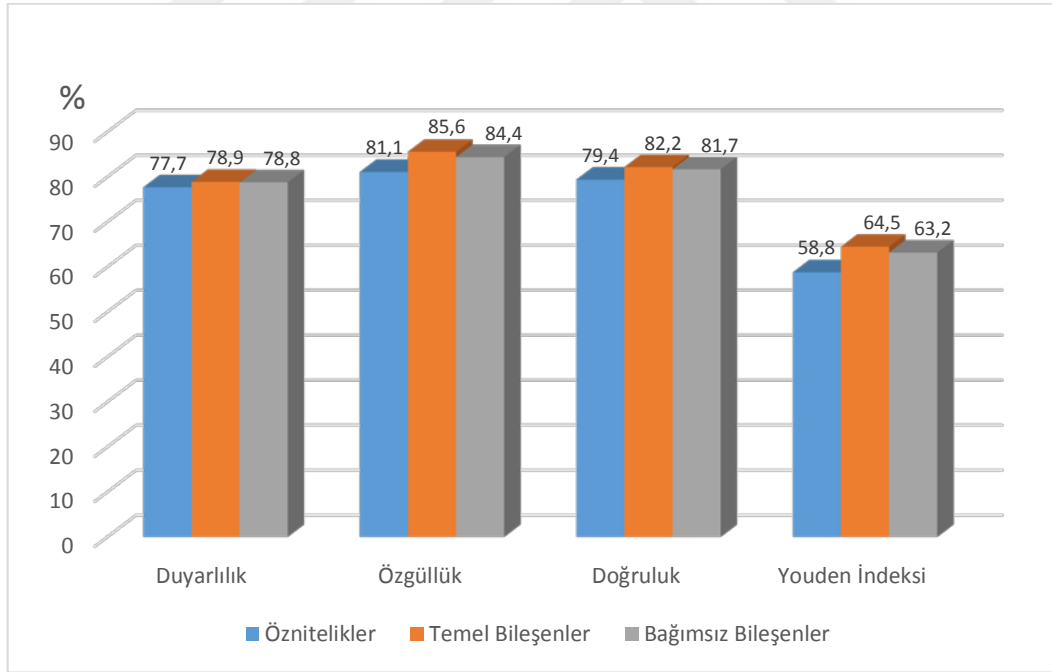
Tablo 4.8 Bağımsız bileşenlere ait Eğitim ve test veri setlerinin Lineer-DVM sınıflandırma sonuçları

	C	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	1	80,0	85,2	82,6	65,2
Test Veri Seti	-	78,8	84,4	81,7	63,2

Lineer-DVM ile yapılan sınıflandırma sonuçlarına göre eğitim veri seti için en yüksek duyarlılık (%91,4), özgüllük (%98,1), doğruluk (%94,7) ve Youden İndeksi (%89,5) öznitelikler için elde edildi (Şekil 4.12). Test veri setinde ise en yüksek duyarlılık (%78,9), özgüllük (%85,6), doğruluk (%82,2) ve Youden İndeksi (%64,5) temel bileşenler için elde edildi (Şekil 4.13).



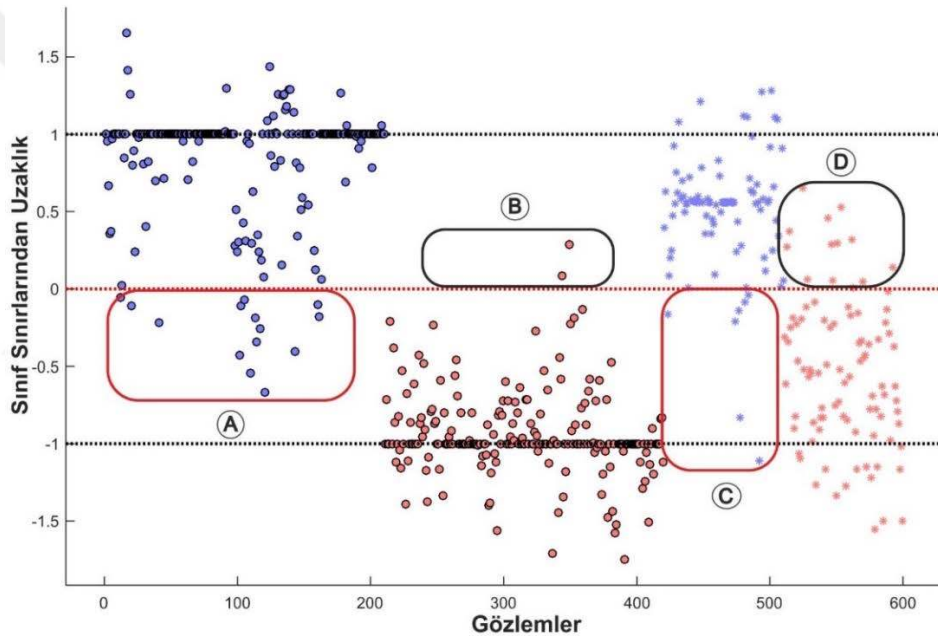
Şekil 4.12 Eğitim veri setlerinin Lineer-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırılması



Şekil 4.13 Test veri setlerinin Lineer-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırılması

4.3.2 Radyal-DVM Sınıflandırma Sonuçları

Özniteliklerle yapılan sınıflandırmada eğitim ve test veri setleri için özniteliklerin sınıf sınırlarından uzaklıkları Şekil 4.14'te gösterildi. Şekil 4.14'te A bölgesine düşen noktalar eğitim veri setinde yanlış sınıflandırılan epileptik sinyalleri ve B bölgesine düşen noktalar ise eğitim veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Benzer şekilde C bölgesine düşen noktalar test veri setinde yanlış sınıflandırılan epileptik sinyalleri ve D bölgesine düşen noktalar ise test veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Şekil 4.14'te test ve eğitim veri setlerinde yanlış sınıflandırılan epileptik noktaların oranlarının benzer olduğu, normal sinyalleri temsil eden noktaların test veri setindeki yanlış sınıflandırma oranının eğitim veri setine göre daha yüksek olduğu görülmektedir.



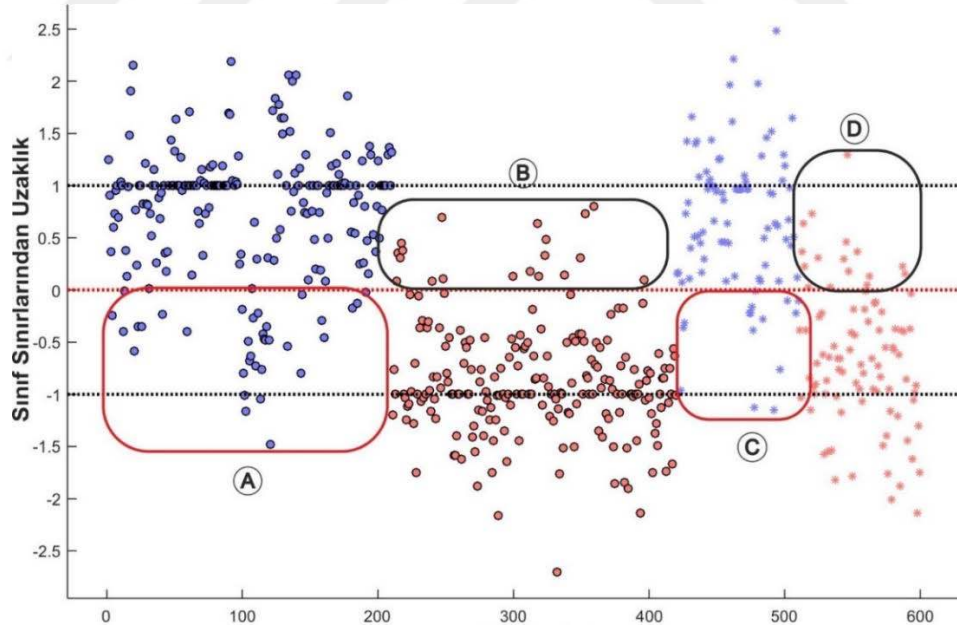
Şekil 4.14 Özniteliklerle yapılan sınıflandırmada eğitim ve test veri setleri için özniteliklerin sınıf sınırlarından uzaklıkları

Eğitim ve test veri setlerinin Radyal-DVM ile oluşturulan sınıflandırma modeline ilişkin ilişkin düzenleme ve genişleme parametreleri ile duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.9'da verildi.

Tablo 4.9 Özniteliklere ait eğitim ve test veri setlerinin Radyal-DVM sınıflandırma sonuçları

	C	γ	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	1	9	93,3	99,1	96,2	92,4
Test Veri Seti	-	-	92,2	85,6	88,9	77,8

Temel bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için temel bileşenlerin sınıf sınırlarından uzaklıkları Şekil 4.15'te gösterildi. Şekil 4.15'te A bölgesine düşen noktalar eğitim veri setinde yanlış sınıflandırılan epileptik sinyalleri ve B bölgesine düşen noktalar ise eğitim veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Benzer şekilde C bölgesine düşen noktalar test veri setinde yanlış sınıflandırılan epileptik sinyalleri ve D bölgesine düşen noktalar ise test veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Şekil 4.15'te test ve eğitim veri setlerinde yanlış sınıflandırılan epileptik noktaların oranlarının benzer olduğu, normal sinyalleri temsil eden noktaların test veri setindeki yanlış sınıflandırma oranının eğitim veri setine göre daha yüksek olduğu görülmektedir.



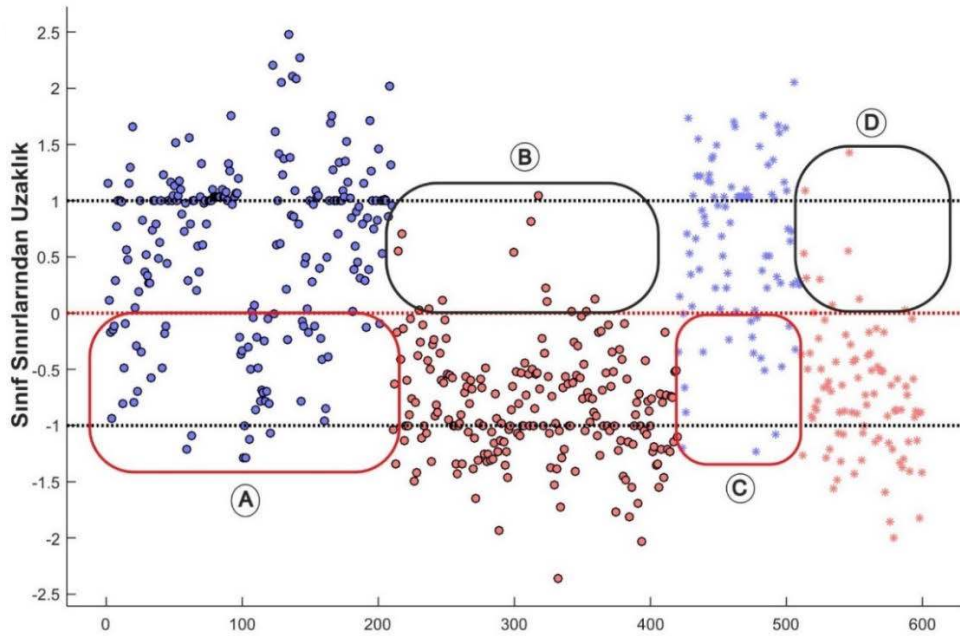
Şekil 4.15 Temel bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için temel bileşenlerin sınıf sınırlarından uzaklıkları

Eğitim ve test veri setlerinin Radyal-DVM ile yapılan sınıflandırma sonuçlarına ilişkin duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.10'da verildi.

Tablo 4.10 Temel bileşenlere ait Eğitim ve test veri setlerinin Radyal-DVM sınıflandırma sonuçları

	C	γ	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	10	6,4	83,8	90,9	87,4	74,7
Test Veri Seti	-	-	83,3	82,2	82,8	65,5

Bağımsız bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için bağımsız bileşenlerin sınıf sınırlarından uzaklıkları Şekil 4.16’da gösterildi. Şekil 4.16’da A bölgesine düşen noktalar eğitim veri setinde yanlış sınıflandırılan epileptik sinyalleri ve B bölgesine düşen noktalar ise eğitim veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Benzer şekilde C bölgesine düşen noktalar test veri setinde yanlış sınıflandırılan epileptik sinyalleri ve D bölgesine düşen noktalar ise test veri seti için yanlış sınıflandırılan normal sinyalleri göstermektedir. Şekil 4.16’da test veri setinde yanlış sınıflandırılan epileptik noktaların oranının eğitim veri setine göre daha düşük olduğu, normal sinyalleri temsil eden noktaların test veri setindeki yanlış sınıflandırma oranının eğitim veri setine göre daha yüksek olduğu görülmektedir.



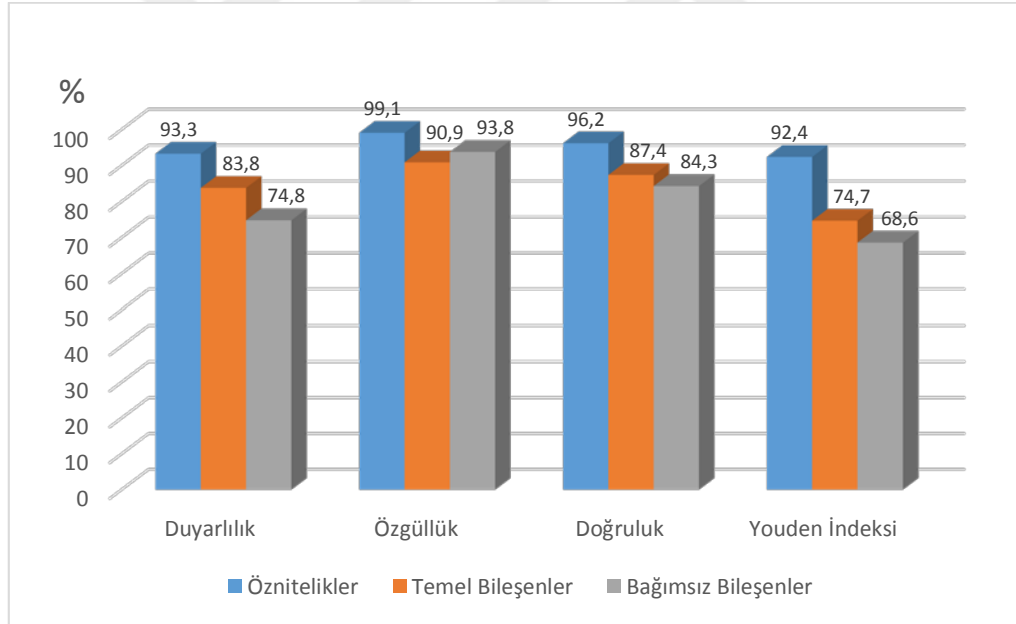
Şekil 4.16 Bağımsız bileşenlerle yapılan sınıflandırmada eğitim ve test veri setleri için gözlemlerin sınıf sınırlarından uzaklıkları

Eğitim ve test veri setlerinin Radyal-DVM ile yapılan sınıflandırma sonuçlarına ilişkin duyarlılık, özgüllük, doğruluk oranları ve Youden İndeksi Tablo 4.11’de verildi.

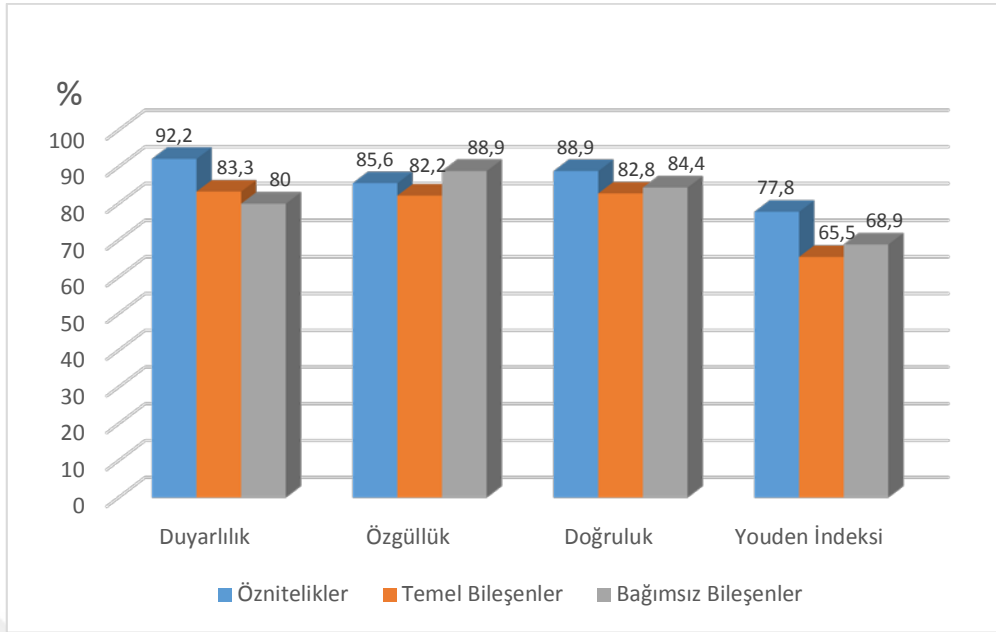
Tablo 4.11 Bağımsız bileşenlere ait eğitim ve test veri setlerinin Radyal-DVM sınıflandırma sonuçları

	C	γ	Duyarlılık (%)	Özgüllük (%)	Doğruluk (%)	Youden İndeksi (%)
Eğitim Veri Seti	10	6,4	74,8	93,8	84,3	68,6
Test Veri Seti	-	-	80,0	88,9	84,4	68,9

Radyal-DVM ile yapılan sınıflandırma sonuçlarına göre eğitim veri seti için en yüksek duyarlılık (%93,3), özgüllük (%99,1), doğruluk (%96,2) ve Youden İndeksi (%92,4) öznitelikler için elde edildi (Şekil 4.17). Test veri setinde ise en yüksek duyarlılık (%92,2), doğruluk (%88,9) ve Youden İndeksi (%77,8) özniteliklerden, en yüksek özgüllük (%88,9) ise bağımsız bileşenlerden elde edildi (Şekil 4.18).

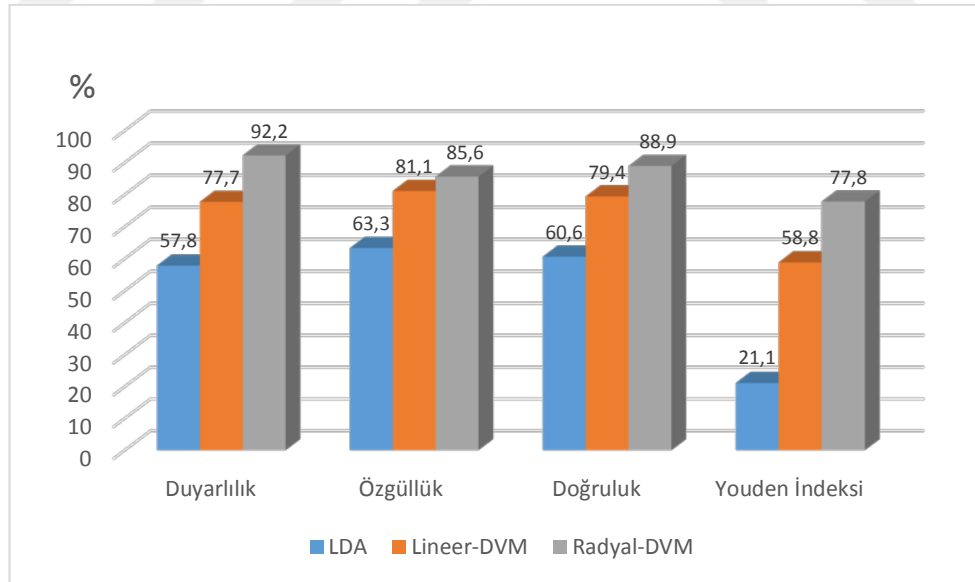


Şekil 4.17 Eğitim veri setlerinin Radyal-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırması



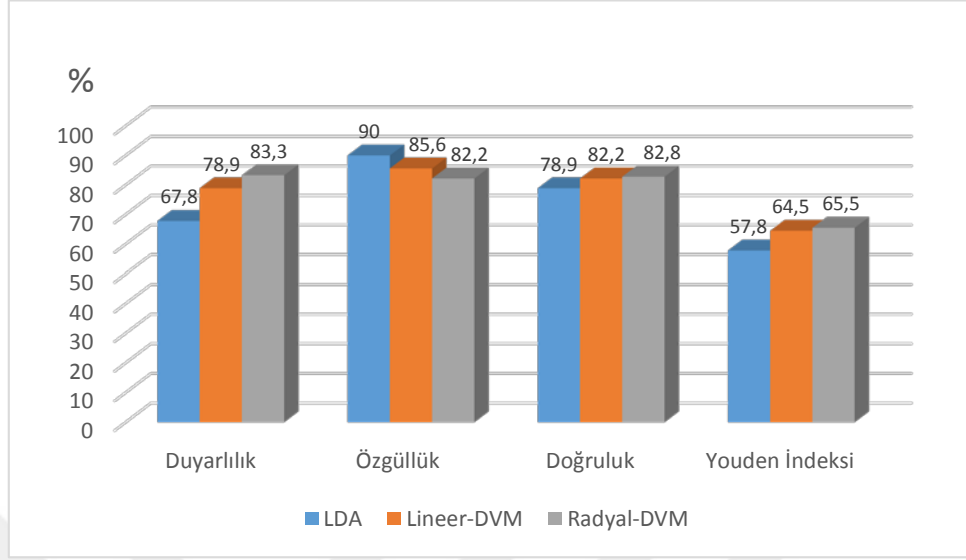
Şekil 4.18 Test veri setlerinin Radyal-DVM ile yapılan sınıflandırma sonuçlarının karşılaştırması

Özniteliklerden sınıflandırma yöntemleri sonucunda elde edilen duyarlılık, özgüllük ve doğruluk oranları Şekil 4.19’da verildi.



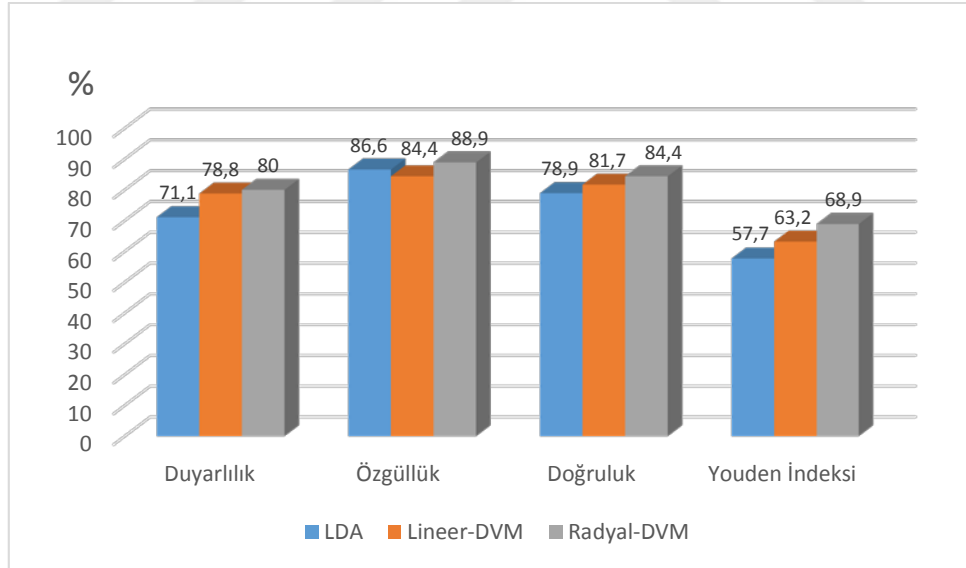
Şekil 4.19 Özniteliklerin sınıflandırma yöntemlerine göre karşılaştırma grafiği

Temel bileşenlerin sınıflandırma yöntemlerine göre sınıflandırma sonuçlarına ilişkin duyarlılık, özgüllük ve doğruluk oranları Şekil 4.20’de verildi.



Şekil 4.20 Temel bileşenlerin sınıflandırma yöntemlerine göre karşılaştırma grafiği

Bağımsız bileşenlerin sınıflandırma yöntemlerine göre sınıflandırma sonuçlarına ilişkin duyarlılık, özgüllük ve doğruluk oranları Şekil 4.21’de verildi.



Şekil 4.21 Bağımsız bileşenlerin sınıflandırma yöntemlerine göre karşılaştırma grafiği

5. TARTIŞMA

Beyin ile ilgili nörolojik bozuklukların (epilepsi, şizofreni, alzheimer vb.) belirlenmesine yönelik yapılan birçok çalışmada EEG sinyalleri kullanılmıştır. EEG sinyallerine dayanan çalışmaların çoğu öznitelik çıkarma ve sınıflandırma olmak üzere iki aşamadan oluşmaktadır. Yapılan çalışmalarda, durağan olmayan bir yapıya sahip olan EEG sinyallerini temsil edecek özniteliklerin çıkarılmasında çeşitli spektral analiz yöntemleri kullanılmıştır. Uygulamalardaki başarısı dikkate alındığında DD yöntemi öne çıkmaktadır (Ocak, 2009). Bu nedenle yapılan birçok çalışmada EEG sinyallerinin spektral analizinde DD tercih edilmiştir.

EEG sinyallerinin spektral analizinde DD'nin kullanıldığı çalışmalardan biri Kıymık ve ark (2004)'nın yaptıkları, EEG sinyallerindeki epileptik sürecin belirlenmesine yönelik çalışmadır. Çalışmada 1,5-3 yaş aralığındaki çocuklardan elde edilen EEG sinyallerinin spektral analizinde KZFD ve DD yöntemlerini kullanmışlar ve epileptik atak belirlemede DD'nin KZFD'ye göre daha başarılı sonuçlar verdiğini göstermişlerdir. Benzer şekilde Coşkun ve İstanbullu (2012) yapmış oldukları çalışmada anestezi altındaki hastanın EEG kayıtlarının spektral analizinde DD ve FD'yi kullanmışlardır. Çalışma sonunda DD'nin FD'ye göre çok daha iyi sonuçlar verdiğini ortaya koymuşlardır. Literatüre bakıldığında, EEG sinyallerinin spektral analizinde DD yönteminin öznitelik çıkarmada oldukça başarılı sonuçlar verdiği ve yaygın bir şekilde kullanıldığı görülmektedir (Subaşı ve Erçelebi, 2005; Lima ve ark, 2010; Subaşı ve Gürsoy, 2010; Wang ve ark, 2013; Rout, 2014). Bu nedenle bu çalışmada, EEG sinyallerinin spektral analizi için DD tercih edildi.

Subaşı ve Erçelebi (2005), 4 kanaldan kaydedilen epileptik olmayan EEG sinyalleri ile uzmanlar tarafından incelenerek seçilen ve farklı epileptik belirtiler içeren EEG sinyallerini ADD yöntemiyle alt frekans bantlarına ayıştırmıştır. α , β , θ ve δ bantlarını temsil eden dalgacık katsayılarından elde ettikleri öznitelikleri, Lojistik Regresyon ile %89,0 ve Yapay Sinir Ağları ile %93,0 doğruluk oranıyla sınıflandırmışlardır. Çalışmamızda kullanılan epileptik sinyaller, Subaşı ve Erçelebi (2005)'nin çalışmasında kullanılan epileptik sinyallere benzer şekilde uzman hekim tarafından incelenerek seçilen ve farklı epileptik belirtiler içeren EEG sinyalleridir. Sinyaller, Subaşı ve Erçelebi (2005)'nin çalışmasına benzer şekilde, ADD ile α , β , θ ve δ alt bantlarına ayıştırıldı ve bu bantları temsil eden dalgacık katsayılarından öznitelikler hesaplandı. DVM ile yapılan sınıflandırmada %88,9 doğruluk oranı elde edildi. Doğruluk oranları arasındaki farklılığın kullanılan sınıflandırma yöntemlerinden çok

alışmalarda kullanılan EEG sinyallerinin ierdiği epileptik aktivitelerin belirginlięi ve oranı ile iliřkili olduęu rlmektedir.

Lima ve ark (2010), hasta olmayan ve epilepsi hastası olan bireylerden kriz anında alınan EEG sinyallerini (Bonn veri tabanı A-E veri setleri) ADD ile alt frekans bantlarına ayırıştırılmıştır. α , β , θ ve δ bantlarını temsil eden dalgacık katsayılarından elde ettikleri znitelikleri DVM ve En Kk Kareler Destek Vektr Makinesi (EKK-DVM) yntemleri ile sınıflandırmış ve en yksek doęruluk oranını EKK-DVM ile elde etmişlerdir. Bir dięer alışmada Orhan ve ark (2011), hasta olmayan ve epilepsi hastalarından kriz anında ve kriz anında olmayan toplam  farklı durumda tek kanaldan kaydedilen EEG sinyallerini (Bonn veri tabanı A-D-E veri setleri) ADD ile alt frekans bantlarına ayırıştırılmış ve bu frekans bantlarından elde ettikleri znitelikleri ok Katmanlı Algılayıcı Yapay Sinir Aęları ile sınıflandırarak %96,6 doęruluk oranı elde etmişlerdir. Kaya ve ark (2012), aynı veri setlerini (Bonn veri tabanı A-D-E veri setleri) ADD yntemiyle alt frekans bantlarına ayırıştırılmış ve alt frekans bantlarından elde ettikleri znitelikleri Karar Kuralları ve Karar Aęaları yntemlerini kullanarak sınıflandırmış ve en yksek doęruluk oranını Karar Aęaları ile elde etmişlerdir. Bir dięer alışmada Rout (2014), aynı veri setlerini (Bonn veri tabanı A-D-E veri setleri) ADD yntemiyle alt frekans bantlarına ayırıştırılmış ve alt frekans bantlarından elde ettięi znitelikleri Lineer-DVM ve AdaBoost DVM yntemlerini kullanarak sınıflandırmış ve en yksek doęruluk oranını AdaBoost DVM ile elde etmişlerdir. alışmamızda epileptik olmayan ve uzman hekim tarafından incelenerek seilen ve farklı epileptik belirtiler ieren EEG sinyallerinden elde edilen znitelikler, LDA, Lineer-DVM ve Radyal-DVM yntemleri kullanılarak sınıflandırıldı ve en yksek doęruluk oranı Radyal-DVM ile elde edildi. Aynı veri setlerinin (Bonn Veri Setleri) kullanıldığı alışmalarda farklı sınıflandırma yntemleriyle elde edilen başarı oranlarının, tercih edilen sınıflandırma yntemlerinin yanı sıra sinyallerden elde edilen ve sınıflandırmada kullanılan zniteliklere baęlı olarak farklılık gsterdięi rlmektedir. Bununla birlikte herhangi bir sınıflandırma ynteminin dięer yntemlerden stn olduęunu net bir řekilde sylemek gtr.

Garrett ve ark (2003), 5 farklı zihinsel durumda 6 kanaldan kaydedilen EEG sinyallerini LDA, YPA ve DVM yntemleri ile sınıflandırmış, sırasıyla %66,0, %69,4 ve %72,0 doęruluk oranı elde etmişlerdir. DVM ile yaptıkları sınıflandırmada alışmamızda olduęu gibi radyal tabanlı ekirdek fonksiyonu kullanmış ve en yksek doęruluk oranını bu yntemle elde etmişlerdir. EEG sinyallerinin sınıflandırılmasında lineer olmayan yntemlerin lineer olan yntemlere gre daha başarılı olduklarını ifade etmişlerdir. Dięer bir alışmada, Lehmann ve ark. (2007), alzeymır hastası olan ve olmayan bireylerden 19 kanal zerinden kaydedilen EEG

sinyallerini kullanarak lineer olan ve olmayan yöntemlerin sınıflandırma performanslarını karşılaştırmışlardır. LDA ile %91,0 ve radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile %95,0 doğruluk oranı elde etmişlerdir. Bu sonuçlara paralel olarak çalışmamızda LDA ile %78,9 ve lineer çekirdek fonksiyonunun kullanıldığı DVM ile %82,2 doğruluk oranı ile sınıflandırma yapılırken radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile %88,9 sınıflandırma başarısı elde edildi.

Subaşı ve Gürsoy (2010) yaptıkları çalışmada, hasta olmayan ve epilepsi hastası olan bireylerden kriz anında tek kanaldan alınan EEG sinyallerini (Bonn veri tabanı A-E veri setleri) ADD ile alt frekans bantlarına ayırtmıştır. α , β , θ ve δ bantlarını temsil eden dalgacık katsayılarından elde ettikleri özniteliklerin boyutlarını LDA, TBA ve BBA yöntemleriyle indirgemiş ve DVM ile sınıflandırmışlardır. LDA ile boyutu indirgenen özniteliklerle %100, BBA ile boyutu indirgenen özniteliklerle %99,5 ve TBA ile boyutu indirgenen özniteliklerle %98,75 doğruluk oranı elde etmişlerdir. Çalışmamızda, hasta olmayan ve epilepsi hastası olan bireylerden kriz geçirmediikleri sırada 16 kanaldan alınan EEG kayıtları, Subaşı ve Gürsoy (2010)'un çalışmasına benzer şekilde, ADD ile α , β , θ ve δ alt bantlarına ayırtıldı ve bu bantları temsil eden dalgacık katsayılarından öznitelikler hesaplandı. Subaşı ve Gürsoy'un çalışmasına paralel olarak çalışmamızda, BBA ile boyut indirgenerek yapılan sınıflandırma sonucu elde edilen doğruluk oranının (%84,4), TBA ile boyut indirgenerek yapılan sınıflandırma sonucu elde edilen doğruluk oranından (%82,8) daha yüksek olduğu gözlemlendi.

Bu çalışmada, hasta olmayan ve epilepsi hastası olan bireylerden kriz geçirmediikleri anda 16 kanal üzerinden kaydedilen EEG sinyallerinden ADD ile toplam 384 adet öznitelik elde edildi. Tüm öznitelikler kullanılarak yapılan sınıflandırmada en yüksek duyarlılık (%92,2), özgüllük (%85,6) ve doğruluk (%88,9) oranı radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile elde edildi. TBA ile elde edilen temel bileşenlerle yapılan sınıflandırmada LDA ve lineer çekirdek fonksiyonunun kullanıldığı DVM yöntemlerinin sınıflandırma performanslarının özniteliklerle yapılan sınıflandırmaya oranla belirgin düzeyde arttığı gözlenirken radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM yönteminin sınıflandırma performansının azaldığı gözlemlendi. Temel bileşenlerle yapılan sınıflandırmalarda en yüksek duyarlılık (%83,3) ve doğruluk (%82,8) oranı radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile en yüksek özgüllük (%85,6) oranı lineer tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile elde edildi. BBA ile elde edilen bağımsız bileşenlerle yapılan sınıflandırmada - temel bileşenlerde olduğu gibi - LDA ve lineer çekirdek fonksiyonunun kullanıldığı DVM yöntemlerinin sınıflandırma performanslarının,

özniteliklerle yapılan sınıflandırmaya oranla belirgin düzeyde arttığı gözlenirken; radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM yönteminin sınıflandırma performansının azaldığı gözlemlendi. Bağımsız bileşenlerle yapılan sınıflandırmalarda en yüksek duyarlılık (%80,0), özgüllük (88,9) ve doğruluk (%84,4) oranı radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile elde edildi. Temel bileşenler ve bağımsız bileşenlerle yapılan sınıflandırmalar kıyaslandığında ise sınıflandırma performanslarının birbirlerine çok yakın olduğu gözlemlendi. Ayrıca temel bileşenler ve bağımsız bileşenlerle yapılan sınıflandırmalarda işlem sürelerinin, özniteliklerle yapılan sınıflandırmaya göre daha kısa olduğu gözlemlendi. Yapılan sınıflandırmalarda hem lineer hem de radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM'nin, LDA'ya göre daha yüksek sınıflandırma başarısına ulaştığı gözlemlendi. Genel anlamda en başarılı sınıflandırma yönteminin ise radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM olduğu belirlendi. Çalışmamızda en yüksek sınıflandırma başarısı özniteliklerin, radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile sınıflandırılması sonucunda elde edildi.

6. SONUÇ ve ÖNERİLER

Çalışmaya uzman hekim tarafından epilepsi tanısı konmuş 10 hasta ve 10 sağlam olmak üzere toplam 20 kişi dahil edildi. Epilepsi tanısı konmuş hastalardan alınan EEG kayıtları uzman hekim tarafından incelenerek seçilen ve farklı epileptik belirtiler içeren kayıtlardı. Epileptik ve epileptik olmayan sinyalleri sınıflandırmak için öncelikle EEG sinyalleri gibi durağan olmayan sinyallerin spektral analizinde oldukça başarılı bir yöntem olan ADD ile sinyaller alt frekans bantlarına ayrıştırıldı ve bu alt frekans bantlarının her birinden literatürde sıkça kullanılmış olan 6 adet istatistik (ortalama, maksimum, minimum, standart sapma, entropi ve enerji) hesaplanarak toplamda 384 adet öznitelik çıkarıldı. Öznitelik uzayının boyutunu azaltmak amacıyla öznitelikler arasındaki korelasyonu ortadan kaldırarak birbirleriyle ilişkisiz temel bileşenler türeten TBA yöntemi ve birbirinden bağımsız bileşenler türeten BBA yöntemi ile özniteliklerin boyutları indirildi. Sınıflandırma, EEG sinyallerinden elde edilen öznitelikler, temel bileşenler ve bağımsız bileşenler kullanılarak üç şekilde gerçekleştirildi. Sınıflandırmalar LDA ve DVM yöntemleri ile yapıldı. DVM ile yapılan sınıflandırmada lineer ve radyal çekirdek fonksiyonları kullanıldı. Böylece hem boyut indirgemedede kullanılan TBA ve BBA yöntemlerinin sınıflandırma başarısı üzerindeki etkileri hem de sınıflandırma için kullanılan LDA, lineer ve radyal tabanlı çekirdek fonksiyonlarının kullanıldığı DVM yöntemlerinin performansları karşılaştırıldı.

Sonuç olarak öznitelikler üzerinde TBA ve BBA yöntemleri ile yapılan boyut indirgeme işleminin LDA ve lineer çekirdek fonksiyonunun kullanıldığı DVM yöntemlerinin sınıflandırma performanslarının özniteliklerle yapılan sınıflandırmaya oranla belirgin düzeyde arttığı gözlenirken; radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM yönteminin sınıflandırma performansının azaldığı gözlemlendi. Temel bileşenler ve bağımsız bileşenlerle yapılan sınıflandırmalar kıyaslandığında ise sınıflandırma performanslarının birbirlerine çok yakın olduğu gözlemlendi. Epileptik ve normal EEG sinyallerinden elde edilen özniteliklerle doğrudan veya boyutu indirgenerek yapılan sınıflandırmalarda radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM yönteminin LDA ve lineer çekirdek fonksiyonunun kullanıldığı DVM yöntemlerine göre hasta ve kontrol bireylerine ait sinyalleri ayırt etme bakımından daha başarılı olduğu belirlendi. En yüksek sınıflandırma başarısı özniteliklerin, radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile sınıflandırılması sonucunda elde edildi. Radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM'nin boyut indirgeme yöntemlerinden olumsuz olarak etkilendiği ve sınıflandırma performansının azaldığı gözlemlendi.

Böylece radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM'nin sınıflandırma performansının, bağımsız değişken sayısından ve bu değişkenler arasındaki ilişkilerden etkilendiği, boyut indirgeme sonucunda veride oluşan bilgi kaybından olumsuz yönde etkilendiği sonucuna varıldı. LDA ve lineer çekirdek fonksiyonunun kullanıldığı DVM'de ise boyut indirgeme yöntemlerinin sınıflandırma performansını arttırdığı sonucuna varıldı.

Uygulama sonuçları gösteriyor ki;

1. ADD katsayıları EEG sinyallerini temsil edebilen katsayılardır ve bu katsayılardan çıkartılan öznitelikler kullanılarak epileptik ve normal sinyaller arasında iyi bir ayrım elde edilebilir.
2. Epileptik ve normal EEG sinyallerinin başarılı bir şekilde sınıflandırılması, başta EEG sinyallerindeki epileptik aktivitelerin oranına ve belirginliğine, sınıflandırmada kullanılan özniteliklere, öznitelikler üzerinde uygulanacak boyut indirgeme yöntemlerine ve sınıflandırmada kullanılacak yöntemlere bağlı olarak değişir.
3. Radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM yöntemi, epilepsi hastalarını sağlam bireylerden ayırt etmede sağladığı yüksek başarıdan dolayı uzman hekimlere yardımcı bir karar destek aracı olarak kullanılabilir.

7. KAYNAKLAR

- Adeli H, Zhou Z, Dadmehr N.** Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of neuroscience methods*, 2003, 123(1), s 69-87.
- Ahammad N, Fathima T, Joseph P.** Detection of epileptic seizure event and onset using EEG. *BioMed research international* 2014, 2014(2014), s 1-7.
- Altunay S, Telatar Z, Erogul O.** Epileptic EEG detection using the linear prediction error energy. *Expert Systems with Applications*, 2010, 37(8), s 5661-5665.
- Andrzejak IA, Ali AA, Abdulridha HM.** Classification of Electroencephalograph (EEG) Signals Using Quantum Neural Network. *Signal Processing: An International Journal (SPIJ)*, 2011, 4 (6), s 329-337.
- Araki S, Makino S, Blin A, Mukai R, Sawada H.** Underdetermined blind separation for speech in real environments with sparseness and ICA. *In Acoustics, Speech, and Signal Processing*, s 881, May 17-21, 2004, Montreal, Quebec, Canada.
- Blankertz B, Curio G, Muller KR.** Classifying single trial EEG: Towards brain computer interfacing. *Advances in neural information processing systems*, 2002, 1, s 157-164
- Bronzino JD.** Biomedical engineering handbook (Vol. 2). CRC press, 1999, 1656.
- Bronzino JD.** Principles of Electroencephalography. The Biomedical Engineering Handbook: Second Edition. Ed. Joseph D. Bronzino Boca Raton: CRC Press LLC, 2000, 3024.
- Burges CJ.** A tutorial on support vector machines for pattern recognition, *Knowledge Discovery and Data Mining*, 1998, 2, s 121-167
- Carter R.** The human brain book. Published in the United States by DK Publishing, 2014. 264.
- Chen G.** Automatic Eeg Seizure Detection Using Dual-Tree Complex Wavelet-Fourier Features. *Expert Systems with Applications*, 2014, 41 (2014), s 2391-2394.
- Cortes C, Vapnik VN.** Support vector networks, *Machine Learning*, 1995, 20(3), s 273-297
- Coşkun M, İstanbullu A.** EEG İşaretlerinin FFT ve Dalgacık Dönüşümü ile Analizi. *XIV. Akademik Bilişim Konferansı*, 1-3 Şubat 2012, Uşak.
- Dunteman GH.** Principal component analysis. Quantitative applications in the social sciences series (vol. 69), 1989, 96.
- Ersöz A, Özşen S.** Uyku EEG Sinyalinin Yapay Sinir Ağ Modeli ile Sınıflandırılması *Elektrik-Elektronik ve Bilgisayar Sempozyumu*, s 298-301, 05-07 Ekim 2011, Elazığ.
- Fielding A.** Cluster and Classification Techniques for the Biosciences, Cambridge University Press, 2007, 260.
- Fodor J, Klempous R, Araujo CPS.** Recent advances in intelligent engineering systems (Vol. 378), Springer Science & Business Media, 2011, 440.
- Gadhok N, Kinsner W.** A study of outliers for robust independent component analysis. *Electrical and Computer Engineering*, 2004. *Canadian Conference on (Volume:3)*, 1421-1425 2-5 May 2004, Canada.

- Gandhi T, Panigrahi BK, Anand S.** A comparative study of wavelet families for EEG signal classification. *Neurocomputing*, 2011, 74(17), 3051-3057.
- Garrett D, Peterson DA, Anderson CW, Thaut MH.** Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. *IEEE Transactions on neural systems and rehabilitation engineering*, 2003, 11(2), s 141-144.
- Geetha G, Geethalakshmi SN.** Artifact Removal from EEG using Spatially Constrained Independent Component Analysis and Wavelet Denoising with Otsu's Thresholding Technique. *Procedia Engineering*, 2012, 30, s 1064-1071.
- Girolami M.** Self-Organising Neural Networks, In *Self-Organising Neural Networks* Springer London, 1999, s 47-75.
- Gonzalez RC, Woods RE.** Digital Image Processing, Second Edition, Prentice-Hall, Inc, New Jersey, 2002, s 779.
- Graps A.** An Introduction to Wavelets, *IEEE Computational Sciences and Engineering*, 1995, Volume 2, Number 2, s 50-61.
- Güler İ, Übeyli ED.** Dalgacık Dönüşümü ile EEG İşaretlerinden Çıkarılan Öznitelik Vektörleri Üzerinde İstatistiksel İşlemlerin Gerçekleştirilmesi. *Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu ELECO*, s 230-234, 2004, Bursa.
- Hardle W, Simar L.** Applied Multivariate Statistical Analysis, First Ed., Springer-Verlag, USA, 2003, 488.
- He P, Wilson G, Russell C.** Removal of Ocular Artifacts From Electroencephalogram by Adaptive Filtering, *Medical and Biological Engineering and Computing*, 2004, s 407-412
- Hyvärinen A, Karhunen J, Erkki O.** Independent Component Analysis, John Wiley & Sons, 2001, 505.
- Hyvärinen A, Oja E.** Independent component analysis: algorithms and applications. *Neural networks*, 2000, 13(4), s 411-430.
- Hyvärinen A.** Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, 1999, 10(3), s 626-634.
- Hyvärinen A.** One-unit contrast functions for independent component analysis: A statistical analysis. In *Neural Networks for Signal Processing VII. Proceedings of the 1997 IEEE Workshop*, 1997, s 388-397
- Johnson RA, Wichern DW.** Applied Multivariate Statistical Analysis, Sixth Ed., Prentice-Hall Inc., USA, 2007, 794.
- Jolliffe IT.** Principal component analysis. John Wiley & Sons, Ltd. 2002, 519.
- Kalaycı Ş.** SPSS uygulamalı çok değişkenli istatistik teknikleri, 2.baskı, Asil Yayın Dağıtım, ANKARA, 2006, 425.
- Kandaswamy A, Kumar CS, Ramanathan RP, Jayaraman S, Malmurugan N.** Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine*, 2004, 34(6), s 523-537
- Kay SM, Marple SL.** Spectrum analysis—a modern perspective. *Proceedings of the IEEE*, 1981, 69(11), 1380-1419.

- Kay SM.** Modern spectral estimation: theory and application. Prentice Hall, Englewood Cliffs, New Jersey, 1988, s 1-286
- Kaya Y, Ertuğrul ÖF, Tekin R.** Epileptik EEG İşaretlerinin Sınıflandırılmasında Karar Kuralları ve Karar Ağaçlarının Kullanılması. *Batman University Journal of Life Sciences*, 2012, 1,(2) : 403-413
- Kıymık MK, Güler I, Dizibüyük A, Akın M.** Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application. *Computers in biology and medicine*, 2005, 35(7), s 603-616
- Klimesch W.** EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 1999, 29(2), s 169-195
- Kumar SP, Sriraam N, Benakop PG, Jinaga BC.** Entropies based detection of epileptic seizures with artificial neural network classifiers. *Expert Systems with Applications*, 2010, 37(4), s 3284-3291.
- Lehmann C, Koenig T, Jelic V, Prichep L, John RE, Wahlund LO, Dierks T.** Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *Journal of neuroscience methods*, 2007, 161(2), s 342-350
- Li Y, Wen P.** Classification of EEG signals using sampling techniques and least square support vector machines. *Rough Sets and Knowledge Technology*. Springer Berlin Heidelberg, 2009, s 375-382
- Lima CA, Coelho AL, Eisencraft M.** Tackling EEG signal classification with least squares support vector machines: a sensitivity analysis study. *Computers in biology and medicine*, 2010, 40(8), s 705-714
- Liu Y, Sourina O, Nguyen KM.** Real-Time Eeg-Based Emotion Recognition And Its Applications. *Transactions on Computational Science*, 2011, s 256-277
- Mahajan K, Vargantwar MR, Rajput SM.** Classification of EEG using PCA, ICA and Neural Network. *International Journal of Engineering and Advanced Technology.(IJEAT)*, 2011, 1(1), s 1-5
- Malmivuo J, Plonsey R.** Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields, Oxford University Press, New York, 1995, 512.
- Masnan MJ, Zakaria A, Shakaff AYM, Mahat NI, Hamid H, Subari, N, Saleh JM.** Principal Component Analysis–A Realization of Classification Success in Multi Sensor Data Fusion. *Principal Component Analysis–Engineering Applications*, In Tech, 2012, s 1-24.
- Merry RJE.** Wavelet Theory And Applications: A Literature Study. Technische Universiteit Eindhoven (Eindhoven), DCT 2005, s 1-41
- Mirowski P, Madhavan D, Lecun Y, Kuzniecky R.** Classification of patterns of EEG synchronization for seizure prediction. *Clinical Neurophysiology*, 2009, 120(11), s 1927-1940
- Misiti M, Misiti Y, Oppenheim G, Poggi J.** Matlab Wavelet Toolbox User's Guide. Version 3. The Mathworks. Inc. Natick, MA., 2004.
- Misiti M, Misiti Y, Oppenheim G, Poggi JM.** Wavelet toolbox user's guide. *The Math Works Ins.*, 1996.

- Misiti M, Misiti Y, Oppenheim G, Poggi JM.** Wavelet toolbox users guide, MathWorks, Inc, 2015.
- Mitchell DJ, McNaughton N, Flanagan D, Kirk IJ.** Frontal-midline theta from the perspective of hippocampal “theta”. *Progress in neurobiology*, 2008, 86(3), s 156-185
- Mohammad N, Achmad J, Mohamad AM.** Automatic Detection of Epileptic Spikes Based on Wavelet Neural Network. *IEEE APCCAS Asia- Pacific Conference on Circuits and Systems*, 2002, 2, s 483–486
- Morlet J, Arens G, Fourgeau E, Giard D.** Wave propagation and sampling theory - part I: complex signal and scattering in multilayered media, *Geophysics*, 1982, s 203-221
- Niedermeyer E, da Silva FL.** Electroencephalography , 5th Ed., Lippincott Williams & Wilkins, USA, 2005, 1309.
- Nunez, PL, Srinivasan R.** *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006, 611.
- Ocak H.** Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with Applications*, 2009, 36(2), s 2027-2036
- Oppenheim AV, Verghese GC.** Signals, systems and inference. Prentice Hall, 2015, 561.
- Oppenheim AV, Willsky AS, Nawab SH.** Signals and systems. Pearson Education Limited, USA, 2014, 937.
- Orhan U, Hekim M, Ozer M.** EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. *Expert Systems with Applications*, 2011, 38(10), s 13475-13481
- Orhan U.** EEG işaretlerinden epilepsi hastalığının teşhisi için yeni yaklaşımlar, Doktora Tezi, Zonguldak Karaelmas Üniversitesi Fen Bilimleri Enstitüsü, Zonguldak 2011, 83.
- Özdamar K.** Paket Programlarla İstatistiksel Veri Analizi 2, 9.Baskı, Nisan Kitabevi, Eskişehir, 2013, 474.
- Özdemir N, Yıldırım E.** Epileptic seizureprediction based on Hilbert Huang Transform and Artificial Neural Networks. In: *Signal Processing and Communications Applications Conference (SIU), 2012 20th*. IEEE, s 1-4, 18-20 Nisan 2012, Muğla.
- Park Y, Luo L, Parhi KK, Netoff T.** Seizure prediction with spectral power of EEG using cost-sensitive support vector machines. *Epilepsia*, 2011, 52(10), s 1761-1770
- Polikar R.** The story of wavelets. *Physics and modern topics in mechanical and electrical engineering*, 1999, s 192-197
- Proakis JG, Manolakis DG.** Digital Signal Processing, Prentice Hall, New Jersey, 2007, s 147-223.
- Rencher AC.** Methods of Multivariate Analysis, Second edition, John Wiley & Sons Inc., USA, 2002, 732.
- Rioul O, Vetterli M.** Wavelets and Signal Processing, *Signal Processing Magazine IEEE*, 1991, 8(4), s 14–38

- Ropper AH, Brown RH.** Adams and Victor's principles of neurology (8th edition). New York: McGraw-Hill Medical. 2005, 1398.
- Ropper AH.** Adams and Victor's principles of neurology (Vol. 9). New York: McGraw-Hill Medical. 2009, 1587.
- Rout N.** Classifications & Misclassifications of EEG Signals using Linear and AdaBoost Support Vector Machines, *International Journal Of Advanced Research, Ideas And Innovations in Technology* 2014, 1(2), s 1-6.
- Sanei S, Chambers JA.** EEG Signal Processing, Wiley, England, 2007, 313.
- Schisterman EF, Perkins NJ, Liu A, Bondell H.** Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology*, 2005, 16, s 73–81
- Schröder P, Sweldens W.** Spherical wavelets: Efficiently representing functions on the sphere. In: *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM, s 161-172, 1995, Los Angeles.
- Semmlow JL.** Biosignal and Medical Image Processing, 2nd edition, CRC Press, 2009, 442.
- Sharma S.** Applied multivariate techniques. John Wiley & Sons, Inc. 1995, 493.
- Stone JV.** Independent Component Analysis: A Tutorial Intoduction. A Bradford Book The MIT Press, 2004, 213.
- Subaşı A, Erçelebi E.** Classification of EEG signals using neural network and logistic regression. *Computer methods and programs in biomedicine*, 2005, 78(2), s 87-99
- Subaşı A.** EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 2007, 32(4), s 1084-1093
- Subaşı, A, Gürsoy MI.** EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications*, 2010, 37(12), s 8659-8666.
- Türe M.** Türkiye’de 1975-1986 Yıllarında Görülen Kimi Bulaşıcı Hastalıkların Fourier Analizi ile Zaman Kalıplarının Belirlenmesi, Yüksek Lisans Tezi, Anadolu Üniversitesi Fen bilimleri Enstitüsü, Eskişehir 1988, 120.
- Ullsperger M, Debener S.** Simultaneous EEG and fMRI: recording, analysis, and application. Oxford University Press. 2010, 311.
- Uyar M, Yildirim S, Gencoglu MT.** An effective wavelet-based feature extraction method for classification of power quality disturbance signals. *Electric Power Systems Research*, 2008, 78(10), s 1747-1755
- Übeyli ED.** Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines. *Computers in Biology and Medicine*, 2008, 38(1), s 14-22
- Übeyli ED.** Combined neural network model employing wavelet coefficients for EEG signals classification. *Digital Signal Processing*, 2009, 19(2), s 297-308.
- Übeyli ED.** Comparison of different classification algorithms in clinical decision-making. *Expert systems*, 2007, 24(1), s 17-31
- Vapnik VN.** The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995, 201.

Vetterli M, Kovacevic J. Wavelets and subband coding. Prentice-hall, 1995, 499.

Wang Y, Zhou W, Yuan Q, Li X, Meng Q, Zhao X, Wang J. Comparison of ictal and interictal EEG signals using fractal features. *International journal of neural systems*, 2013, 23(06), 1350028.

Web_1. (2009) <http://www.coheadquarters.com/coOuterBrain1x.jpg>, (04.04.2016)

Xizheng Z, Ling Y, Weixiong W. Wavelet Time-frequency Analysis of Electroencephalogram (EEG) Processing, *International Journal of Advanced Computer Science and Applications*, 2010, 1(5), s 1-5.

Yazgan E, Korürek M. Tıp Elektronöğü, İTÜ, İstanbul, 1996, 300.

Ye Y, Zhang ZL, Zeng J, Peng L. A Fast and Adaptive Ica Algorithm with Its Application to Fetal Electrocardiogram Extraction. *Applied Mathematics and Computation* ,2008, 205, 799-806.

Yuan Y, Li Y, Yu L, Guo H. Analysis of nonlinearity in normal and epileptic EEG signals. *In Signal Processing, 2008. ICSP 2008. 9th International Conference on IEEE*, s 2162-2165 2008, Beijing, China.

8. EKLER

Ek 1

Evrak Tarih ve Sayısı: 07/04/2016-E.13777



T.C.
ADNAN MENDERES ÜNİVERSİTESİ REKTÖRLÜĞÜ
Tıp Fakültesi Dekanlığı
Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu



Sayı : 53043469-050.04.04
Konu : Etik Kurul Kararları

Sayın Prof.Dr. Mevlüt TÜRE
Anabilim Dalı Başkanı

Adnan Menderes Üniversitesi Tıp Fakültesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulunun 31.03.2016 tarihinde yapılan olağan toplantısında çalışmanızı ilgilili alınan 17 nolu karar aşağıda sunulmuştur.
Bilgilerinize sunarım

e-imzalıdır
Prof.Dr. Nefati KIYLIOĞLU
Kurul Başkanı

KARAR 17

Protokol No : 2016/873

Sorumlu Yürütücü : Prof. Dr. Mevlüt TÜRE
Biyoistatistik AD

Tıp Fakültesi Biyoistatistik Anabilim Dalı Öğretim Üyesi Prof.Dr. Mevlüt TÜRE'nin "**EEG sinyallerinde farklı boyut indirgeme ve sınıflandırma yöntemlerinin karşılaştırılması**" konulu yukarıda bilgileri verilen klinik araştırma başvuru dosyası ile ilgili belgeler araştırmanın gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş ve uygun bulunmuş olup, çalışmanın başvuru dosyasında belirtilen merkezlerde gerçekleştirilmesinde etik ve bilimsel sakınca bulunmadığına oy birliğiyle karar verilmiştir.

Yine sorumlu araştırmacıya; Form 2'nin 14.1.'in son bölümünde taahhüt edilen **çalışma bittikten sonra nihai raporun, [Sonuç Raporu (web'te), BGOF (Bilgilendirilmiş Gönüllü Olur Formu- gönüllüler tarafından bizzat kendilerinin kendi adı-soyadını yazması ve imzalamasının sağlanması ile adreslerinin eksiksiz olarak formlara yazılmasına dikkat edilmelidir.) ve ORF (Olgu Rapor Formu/Anket)]lerin gönderilmesi gerektiğinin hatırlatılmasına** ve sorumlu yürütücülerinin bu hususa özen göstermesi gerektiğinin bir kez daha vurgulanmasına oy birliğiyle karar verilmiştir.

Sorumlu yürütücü Prof.Dr. Mevlüt TÜRE Etik kurul Üyesi ve yardımcı araştırmacı Prof.Dr.Nefati KIYLIOĞLU da Etik Kurul Başkanı olduğundan dolayı oylamaya katılmamıştır

ÖZGEÇMİŞ

Soyadı, Adı : ÖZTÜRK, Hakan
Uyruk : T.C
Doğum yeri ve tarihi : Konak / İZMİR, 18.05.1983
Telefon : 0256 213 78 39 - 4535
E-mail : hakan.ozturk@adu.edu.tr
Yabancı Dil : İngilizce (Konuşma: İyi, Yazma : İyi, Okuma : İyi)
(ÜDS=80)

EĞİTİM

Derece	Kurum	Bulunduğu Yıllar
Y. Lisans	Abant İzzet Baysal Üniversitesi, Matematik Bölümü	2006 - 2009
Lisans	Abant İzzet Baysal Üniversitesi, Matematik Bölümü	2001 - 2006
Lise	İzmir Atatürk Lisesi	1997 - 2001

İŞ DENEYİMİ

Yıl	Yer/Kurum	Ünvan
2014 -	Adnan Menderes Üniversitesi Biyoistatistik Anabilim Dalı	Araştırma Görevlisi
2011 - 2014	İBB Sosyal Projeler Müdürlüğü	Eğitmen
2007 - 2014	Özel Sektör	Matematik Öğretmeni

KATILDIĞI KURS, KONGRE VE TOPLANTILAR

1. 16. Ulusal Biyoistatistik Kongresi, 10-12 Eylül 2014, Side-Antalya
2. 17. Ulusal Biyoistatistik Kongresi, 5-9. Kasım. 2015, Girne-KKTC

AKADEMİK YAYINLAR

Emre Ayıntap, Ersin Yücel, Bediz Özen, **Hakan Öztürk**, Burcu Işık, Giray Ersöz, “Konjetinal makrotrombositopenili bir olguda presenil katarakt”, *Türk Oftalmoloji Derneği 49. Ulusal Kongresi*, PS0132, İstanbul, 4-8. Kasım. 2015.

Bediz Özen, Gamze Türe, **Hakan Öztürk**, Şefika Kolağası, Giray Ersöz, “Hidroklorotiyazid kullanımı ve akut açılı kapanması glokomu: Bir olgu sunumu”, *Türk Oftalmoloji Derneği 49. Ulusal Kongresi*, PS0077, İstanbul, 4-8. Kasım. 2015.