

**FATİH SULTAN MEHMET VAKIF ÜNİVERSİTESİ**  
**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**ÖZİNİTELİK ÇIKARIM VE EVRİMSEL**  
**ÖZİNİTELİK SEÇİM METOTLARININ**  
**EEG SİNYALLERİNİN SINIFLANDIRMA**  
**BAŞARISINA ETKİLERİ**

**YÜKSEK LİSANS TEZİ**

**Ferda ABBASOĞLU**

**Bilgisayar Mühendisliği Anabilim Dalı**  
**Bilgisayar Mühendisliği Tezli Yüksek Lisans**

**OCAK 2019**



**FATİH SULTAN MEHMET VAKIF ÜNİVERSİTESİ  
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**ÖZNİTELİK ÇIKARIM VE EVRİMSEL  
ÖZNİTELİK SEÇİM METOTLARININ  
EEG SİNYALLERİNİN SINIFLANDIRMA  
BAŞARISINA ETKİLERİ**

**YÜKSEK LİSANS TEZİ**

**Ferda ABBASOĞLU**

**(160221003)**

**Anabilim Dalı: Bilgisayar Mühendisliği**

**Tez Danışmanı:**

**Dr. Öğr. Üyesi Ayla GÜLCÜ**

**Dr. Öğr. Üyesi Ulvi BAŞPINAR (Eş Danışman)**

**Teslim Tarihi: 14.12.2018**

FSMVÜ, Lisansüstü Eğitim Enstitüsü'nün Bilgisayar Mühendisliği Tezli Yüksek Lisans öğrencisi 160221003 numaralı Ferda ABBASOĞLU, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “Öz nitelik Çıkarım ve Evrimsel Öz nitelik Seçim Metotlarının EEG Sinyallerinin Sınıflandırma Başarısına Etkileri” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

**Tez Danışmanı : Dr. Öğr. Üyesi Ayla Gülcü** .....

**Fatih Sultan Mehmet Vakıf Üniversitesi**

**İkinci Tez Danışmanı : Dr. Öğr. Üyesi Ulvi BAŞPINAR** .....

**Marmara Üniversitesi**

**Jüri Üyeleri : Prof. Dr. A. Yılmaz ÇAMURCU** .....

**Fatih Sultan Mehmet Vakıf Üniversitesi**

**Doç. Dr. Üyesi Turgay Tugay BİLGİN** .....

**Bursa Teknik Üniversitesi**

**Dr. Öğr. Üyesi Ayla Gülcü** .....

**Fatih Sultan Mehmet Vakıf Üniversitesi**

**Teslim Tarihi : 14 Aralık 2018**

**Savunma Tarihi : 14 Ocak 2019**

## ÖNSÖZ

Bu tez Fatih Sultan Mehmet Vakıf Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Programı'nda hazırlanmıştır.

Tez çalışması süresince beni destekleyen ve benden yardımlarını esirgemeyen değerli danışman hocalarım Dr. Öğr. Üyesi Ayla GÜLCÜ ve Dr. Öğr. Üyesi Ulvi BAŞPINAR'a teşekkürü bir borç bilirim. Hayatım boyunca benden maddi manevi desteklerini hiç esirgemeyen anneme, babama ve abime teşekkür ederim. Bu tezin bundan sonraki çalışmalara katkı sağlamasını temenni ederim.

OCAK 2019

Ferda ABBASOĞLU

## İÇİNDEKİLER

ÖNSÖZ.....	iv
KISALTMALAR LİSTESİ .....	vii
ÇİZELGE LİSTESİ.....	viii
ŞEKİL LİSTESİ.....	x
SEMBOL LİSTESİ.....	xi
ÖZET.....	xii
SUMMARY.....	xiv
1. GİRİŞ .....	1
1.1. EEG Nedir?.....	2
1.1.1.Teta Dalgaları: .....	3
1.1.2. Delta Dalgaları: .....	4
1.1.3. Alfa Dalgaları: .....	4
1.1.4. Beta Dalgaları: .....	4
1.1.5. Gama Dalgaları .....	4
1.2. Kaynak Araştırması .....	4
1.2.1. Bonn Üniversitesinin Epilepsi Verisiyle Yapılmış Çalışmalar.....	5
1.2.2. BCI Yarışma Verisiyle Yapılmış Çalışmalar .....	7
1.2.3. Diğer Veri Setleri ile Yapılmış Çalışmalar .....	8
2. VERİ İŞLEME VE ÖZNETELİK ÇIKARIMI .....	10
2.1. Kullanılan Verilerin Tanımlanması .....	10
2.1.1. Bonn Üniversitesi Verisi.....	10
2.1.2. BCI Yarışma Verisi .....	12
2.1.3. Ön işleme ve Gürültü Temizleme .....	15
2.2. Öznitelik Çıkarma ve Boyut İndirgeme.....	19
2.2.1. Ortak Uzamsal Örüntüler (CSP).....	19
2.2.2. Ayrık Dalgacık Dönüşümü (ADD).....	22
2.2.3. Yerel İkili Örüntü (YİÖ).....	25
2.2.4. Temel Bileşen Analizi (TBA) .....	28
2.2.5. Ki Kare Testi ile öznitelik seçimi .....	29
3. EEG VERİSİNİN SINIFLANDIRILMASI .....	31
3.1. Gözetimli Öğretim Yöntemi .....	31
3.1.1. Destek Vektör Makineleri (DVM) .....	31
3.1.1.1. Doğrusal destek vektör makineleri .....	32
3.1.1.2. Doğrusal Olmayan DVM .....	33
3.1.1.3. Çok Sınıflı DVM .....	34
3.1.2. Doğrusal Diskriminant Analiz .....	34
3.2. Diferansiyel Evrim Algoritması (DEA).....	35
3.2.1. Popülasyon Yapısı .....	36
3.2.2 Mutasyon .....	37
3.2.3. Çaprazlama .....	37

3.2.4. Seçim.....	38
3.2.5. Diferansiyel Evrim ile Öznitelik Seçimi .....	38
4. BULGULAR .....	42
4.1. Bonn Verisi İçin Bulgular.....	42
4.1.1. LDA ile Sınıflandırma Doğruluk Oranı .....	45
4.1.1.1. ADD Kullanılarak Öznitelik Matrisi Oluşturma.....	45
4.1.1.2. DEÖS Kullanılarak Öznitelik Oluşturma.....	48
4.1.1.3. TBA Kullanılarak Öznitelik Çıkarımı .....	50
4.1.1.4. 1D-YİÖ Kullanılarak Öznitelik Oluşturma .....	51
4.1.2. DVM ile Sınıflandırma Doğruluk Oranı .....	53
4.1.2.1. ADD Kullanılarak Öznitelik Oluşturma.....	53
4.1.2.2. DEÖS Kullanılarak Öznitelik Oluşturma.....	56
4.1.2.3. 1D-YİÖ Kullanılarak Öznitelik Oluşturma .....	58
4.1.3. Bütün Yöntemlerin Karşılaştırılması .....	60
4.2. BCI Verisi İçin Bulgular .....	61
4.2.1. CSP Uygulanması.....	61
4.2.2. ADD Kullanılarak Öznitelik Oluşturma .....	61
4.2.3. DEÖS Kullanılarak Öznitelik Oluşturma .....	62
4.2.4. Ki Kare ile Öznitelik Oluşturma .....	62
4.2.5. 1D-YİÖ İle Öznitelik Oluşturma .....	62
4.2.6. Bütün Yöntemlerin Karşılaştırılması .....	62
5. SONUÇLAR VE ÖNERİLER.....	64
KAYNAKLAR.....	66
ÖZGEÇMİŞ .....	70

## KISALTMALAR LİSTESİ

<b>Kısaltma</b>	<b>Açıklama</b>
EEG	Elektroensefalografi
Hz	Hertz
BBA	Beyin Bilgisayar Arayüzü
CSP	Ortak Uzamsal Örüntüler (Common Spatial Patterns)
BKK	Bire Karşı Kalanı
ADD	Ayrık Dalgacık Dönüşümü
LDA	Lineer Diskriminant Analizi
DVM	Destek Vektör Makinesi
YSA	Yapay Sinir Ağları
AR	Otoregresif model
DPD	Dalgacık Paket Dönüşümü
SampEn	Örnek Entropisi (Sample entropy)
TBA	Temel Bileşen Analizi
mRMR	Minimum Gereksizlik Maksimum İlgili (Minimum Redundancy Maksimum Relevance)
BK	Bilgi Kazanımı (Information Gain)
KKA	Karıncı Kolonisi Algoritması
DEÖS	Diferansiyel Evrim Öznitelik Seçimi
YİÖ	Yerel İkili Örüntü

## ÇİZELGE LİSTESİ

TABLO 2.1: BONN VERİSİ ÖZETİ . . . . .	11
TABLO 2.2: TETİKLENME BAŞLAMA ÖRNEĞİ . . . . .	13
TABLO 2.3: TETİKLEYİCİ İŞLEM ÖRNEĞİ . . . . .	14
TABLO 2.4: DENEKLER İÇİN EĞİTİM VE TEST SAYILARI . . . . .	15
TABLO 2.5: LİTERATÜRDE BENZER FİLTRELEME ÖRNEKLERİNİN GÖRÜLDÜĞÜ ÇALIŞMALAR. . . . .	17
TABLO 2.6: DALGACIK KATSAYILARININ TEMSİL ETTİĞİ DALGALAR. . . . .	23
TABLO 4.1: LİTERATÜRDE BONN VERİSİNİN FARKLI KÜMELERİ İLE YAPILAN ÇALIŞMALAR VE SINIFLANDIRMA DOĞRULUK ORANLARI . . . . .	44
TABLO 4.2: İKİLİ OLARAK A SINIFI İLE DİĞER SINIFLARIN VARLIĞINDA LDA İLE DOĞRULUK ORANI KİYASLANMASI. . . . .	46
TABLO 4.3: İKİLİ OLARAK B SINIFI İLE DİĞER SINIFLARIN VARLIĞINDA LDA İLE DOĞRULUK ORANI KİYASLANMASI. . . . .	46
TABLO 4.4: İKİLİ OLARAK C VE D SINIFLARI İLE DİĞER SINIFLARIN VARLIĞINDA LDA İLE DOĞRULUK ORANI KİYASLANMASI. . . . .	46
TABLO 4.5: ÜÇLÜ KOMBİNLER HALİNDE SINIFLARIN VARLIĞINDA LDA İLE DOĞRULUK ORANI KİYASLANMASI. . . . .	47
TABLO 4.6: FARKLI SINIFLAR ARASINDA LDA İLE DOĞRULUK ORANI. . . . .	48
TABLO 4.7: 10 ÖZİNİTELİĞE SAHİP ALT KÜMENİN LDA İLE SINIFLANDIRMA DOĞRULUK ORANI GÖSTERİMİ . . . . .	49
TABLO 4.8: 15 ÖZİNİTELİĞE SAHİP ALT KÜMENİN LDA İLE SINIFLANDIRMA DOĞRULUK ORANI GÖSTERİMİ . . . . .	49
TABLO 4.9: 20 ÖZİNİTELİĞE SAHİP ALT KÜMENİN LDA İLE SINIFLANDIRMA DOĞRULUK ORANI GÖSTERİMİ . . . . .	50
TABLO 4.10: TBA İLE SINIFLANDIRMA DOĞRULUK ORANLARININ KİYASLANMASI. . . . .	50
TABLO 4.11: A SINIFI İLE DİĞER SINIFLARIN İKİLİ OLUŞTURDUKLARI KÜMENİN LDA İLE DOĞRULUK ORANI. . . . .	51
TABLO 4.12: B SINIFI İLE DİĞER SINIFLARIN İKİLİ OLUŞTURDUKLARI KÜMENİN LDA İLE DOĞRULUK ORANI. . . . .	51
TABLO 4.13: C VE D SINIFLARININ DİĞER SINIFLAR VARLIĞINDA AYRI AYRI İKİLİ KOMBİNLER İLE OLUŞTURDUKLARI KÜMENİN LDA İLE DOĞRULUK ORANI . . . . .	52
TABLO 4.14: ÜÇLÜ KOMBİNLER HALİNDE KÜMELERİN SINIFLANDIRMA DOĞRULUK ORANLARININ LDA İLE ÖLÇÜLMESİ. . . . .	52
TABLO 4.15: DÖRTLÜ KOMBİNLERİN LDA İLE DOĞRULUK ORANLARININ KİYASLANMASI . . . . .	53
TABLO 4.16: İKİLİ OLARAK A SINIFI İLE DİĞER SINIFLARIN VARLIĞINDA DVM İLE DOĞRULUK ORANLARI KİYASLANMASI. . . . .	53
TABLO 4.17: İKİLİ OLARAK B SINIFI İLE DİĞER SINIFLARIN VARLIĞINDA DVM İLE DOĞRULUK ORANI KİYASLANMASI. . . . .	54
TABLO 4.18: İKİLİ OLARAK C VE D SINIFLARI İLE DİĞER SINIFLARIN VARLIĞINDA DVM İLE DOĞRULUK ORANI KİYASLANMASI. . . . .	54
TABLO 4.19: ÜÇLÜ KOMBİNLER HALİNDEKİ SINIFLARIN VARLIĞINDA DOĞRULUK ORANI KİYASLANMASI. . . . .	55
TABLO 4.20: FARKLI SINIFLAR ARASINDA DVM İLE SINIFLANDIRMA DOĞRULUK ORANLARININ ÖLÇÜLMESİ. . . . .	55



TABLO 4.21 : 10 ÖZİNİTELİĞE SAHİP ALT KÜMENİN DVM İLE SINIFLANDIRMA DOĞRULUK ORANI GÖSTERİMİ .....	56
TABLO 4.22: 15 ÖZİNİTELİĞE SAHİP ALT KÜMENİN DVM İLE DOĞRULUK ORANI GÖSTERİMİ .....	57
TABLO 4.23: 20 ÖZİNİTELİĞE SAHİP ALT KÜMENİN DVM İLE DOĞRULUK ORANI GÖSTERİMİ .....	57
TABLO 4.24: A SINIFI İLE DİĞER SINIFLARIN İKİLİ OLUŞTURDUKLARI KÜMENİN DVM İLE DOĞRULUK ORANI.....	58
TABLO 4.25: B SINIFI İLE DİĞER SINIFLARIN İKİLİ OLUŞTURDUKLARI KÜMENİN DVM İLE DOĞRULUK ORANLARI.....	58
TABLO 4.26: C VE D SINIFLARININ DİĞER SINIFLAR VARLIĞINDA AYRI AYRI İKİLİ KOMBİNLER İLE OLUŞTURDUKLARI KÜMELERİN DVM İLE DOĞRULUK ORANLARI..	59
TABLO 4.27: ÜÇLÜ KOMBİNLERİN DVM İLE DOĞRULUK ORANLARININ KIYASLANMASI	59
TABLO 4.28: AXE VE BXE İÇİN ADD VE 1D-YİÖ DOĞRULUK ORANLARI KARŞILAŞTIRILMASI .....	60
TABLO 4.29: BEŞLİ KÜME İÇİN KULLANILAN BÜTÜN YÖNTEMLERİN KIYASLANMASI. ...	60
TABLO 4.30: BCI VERİSİ İÇİN DENENEN METOTLAR VE DOĞRULUK ORANLARI .....	63
TABLO 4.31: BCI VERİSİ DEÖS ‘ÜN FARKLI CR DEĞERLERİNE GÖRE DOĞRULUK ORANI KIYASLAMASI .....	63

## ŞEKİL LİSTESİ

ŞEKİL 1.1: EEG ÖLÇÜMÜ ÖRNEĞİ .....	2
ŞEKİL 1.2: ALFA BETA, TETA, DELTA BEYİN DALGALARINA AİT FREKANSLAR.....	3
ŞEKİL 2.1: TEZDE YAPILAN İŞLEMLER.....	10
ŞEKİL 2.2: DENEYİN GERÇEKLEŞME AŞAMALARI .....	13
ŞEKİL 2.3: ELEKTROTLARIN KAFADAKİ POZİSYONLARI .....	15
ŞEKİL 2.4: BONN VERİSİ İÇİN UYGULANAN İŞLEMLER .....	16
ŞEKİL 2.5: BCI VERİSİ İÇİN UYGULANAN İŞLEMLER.....	17
ŞEKİL 2.6: ORJİNAL VE FİLTRELENMİŞ SİNYALİN FREKANS ARALIĞI. ....	18
ŞEKİL 2.7: ORJİNAL VE FİLTRELENMİŞ SİNYALİN BANT GEÇİŞLİLİĞİNİN TOPLAM GÜCÜ. .....	18
ŞEKİL 2.8: BANT GEÇİREN FİLTRE UYGULANDIKTAN SONRA VERİNİN FREKANS ARALIĞI. .....	19
ŞEKİL 2.9: EEG 5. SEVİYE DALGACIK AYRIŞMASI.....	23
ŞEKİL 2.10: ADD İLE ÖZİNİTELİK ÇIKARIMI .....	24
ŞEKİL 2.11: MERKEZ PİKSELİN KOMŞU PİKSELLERLE DEĞERİNİN KIYASLANMASI .....	25
ŞEKİL 2.12: ELDE EDİLEN 8 BİTLİK SAYININ 10'LUK TABANA ÇEVİRİLMESİ .....	25
ŞEKİL 2.13: MERKEZ SİNYAL VE KAÇ KOMŞU SİNYALİN OLACAĞINA KARAR VERME ....	26
ŞEKİL 2.14: 1D SİNYALE YİÖ UYGULANMASI .....	27
ŞEKİL 2.15: 256 ARALIĞINDAN 59 ARALIĞINA DÜŞÜRÜLMESİ .....	28
ŞEKİL 3.1: DOĞRUSAL OLARAK AYRILABİLEN VERİLER İÇİN HİPER DÜZLEM VE DESTEK VEKTÖRLERİ ..	33
ŞEKİL 3.2: DİFERANSİYEL EVRİM ALGORİTMASI AKIŞ ŞEMASI .....	36
ŞEKİL 3.3: DEÖS'DE DEĞİŞKENLERİN TANIMLANMASI .....	38
ŞEKİL 3.4: DEÖS'ÜN POPÜLASYON VE HATA ORANININ BULUNMASI.....	39
ŞEKİL 3.5: DAĞILIM FAKTÖRÜNÜN HESAPLANMASI .....	41

## SEMBOL LİSTESİ

$\alpha$	: Alfa
$\beta$	: Beta
$\theta$	: Teta
$\delta$	: Delta
$\chi^2$	: Ki-Kare
$X_1$ :	: Uzamsal kovaryans
$R_0$	: Öz vektör matrisi
$\lambda_0$	: Öz değer matrisi
$P$	: Beyazlatma Matrisi
$W$	: Projeksiyon Matrisi
$DWT()$	: Dalgacık Dönüşümü
$\psi$	: Ana dalgacık
$G(xi)$	: Merkezi Piksel

# ÖZİNİTELİK ÇIKARIM VE EVRİMSEL ÖZİNİTELİK SEÇİM METOTLARININ EEG SİNYALLERİNİN SINIFLANDIRMA BAŞARISINA ETKİLERİ

## ÖZET

EEG sinyalleri beyindeki elektriksel aktivitenin ölçülmesi ile elde edilir. Bu sinyallerin analizi, incelenmesi ve sınıflandırılması epilepsi ve uyku bozukluğu gibi rahatsızlıkların teşhisinde kullanılır. EEG sinyallerinin sınıflandırılarak bir sonuca ulaşılması ve bu sonuçlara ilişkin mantıklı yaklaşımlar getirilmesi bazen güç ve karmaşık bir işlem olabilmektedir. Bu tezde EEG sinyalinin sınıflandırma doğruluk oranını arttırmak için örüntü tanıma tabanlı, öznelik çıkarım ve seçme yöntemleri ile çalışmalar yapılmıştır. Bu tezde ilk olarak kullanılan verilere sınıflandırma doğruluk oranının artırılması amacıyla gürültülü verinin elenmesi için bant geçiren filtre uygulanmıştır. Filtreleme işleminden sonra verilere öznelik çıkarım metotları uygulanmıştır. Öznelik çıkarım işleminden sonra öznelik seçim işlemi ile sınıflandırma doğruluğunu arttıran öznelikler seçilmiştir. Elde edilen öznelik matrisi sınıflandırma metotlarından Destek Vektör Makineleri (DVM) ve Lineer Diskriminant Analiz (LDA) ile test edilmiştir. Sınıflandırma doğruluk oranları karşılaştırılmıştır. Belirtilen yöntemler iki farklı veri üzerinde uygulanmıştır. Bunlardan ilki Bonn Üniversitesi tarafından epilepsi rahatsızlığının tespiti için toplanan veridir. Bu veri kümesi tek kanallı 100 örneğe sahiptir ve 5 farklı sınıftan oluşmaktadır. Bu sınıfların 2'si sağlıklı, 2'si hasta ve 1'i nöbet halindeki kişilerin verileridir. Diğer de Berlin Teknik Üniversitesi tarafından düşünce yolu ile kasların oynatılması üzerine çalışmalar yapılması için toplanan BCI yarışma verisidir. BCI'nin 3. yarışmasının 3a veri kümesi kullanılmıştır. Bu veri kümesi ise 60 kanala sahiptir ve 4 farklı sınıftan oluşmaktadır. Bu sınıflar sağ el, sol el, dil ve ayak hareketlerinin verilerini temsil etmektedir. Öznelik çıkarım işlemi var olan veri kümesinden farklı filtrelemeler ve istatistiksel yöntemleri kullanılarak verinin sınıflandırma işlemine uygun hale getirilmesidir. Bu işlem sırasında sütun yapıları değişmekte ve veriyi temsil edecek yeni bir küme oluşmaktadır. Öznelik çıkarım metotlarından Ortak Uzamsal Örüntüler (CSP), 1D Yerel İkili Örüntü (1D-YİÖ), Ayrık Dalgacık Dönüşümü (ADD) ve Temel Bileşen Analizi (TBA) kullanılmıştır. Öznelik çıkarım yöntemlerinden CSP, ayrı sınıfları ayırt etmek ve elde edilen sinyallerin uzamsal çözünürlüğünü arttırmak için sıkça kullanılan bir uzamsal filtreleme yöntemidir. CSP iki sınıflı kümeler için kullanılmaktadır fakat bu tezde çok sınıflı kümeler için genişletilmiş hali kullanılmıştır. 1D-YİÖ ise sinyal işleme yöntemlerinin birçoğu gibi görüntü işleme yöntemlerinden türetilmiş bir metottur. Sinyali bir görüntü gibi ele alıp merkezi

sinyalin komşu sinyallerden büyük olup olmamasına göre veriyi düzenleyerek istatistiksel özellikler çıkarmaktadır. ADD ise sinyalin farklı alt bantlara bölünmesi ile istenilen frekans aralığındaki verilerin filtrelenerek ortaya çıkmasını ve o aralıktaki verilerin katsayı bilgileri kullanılarak istatistiksel özelliklerin çıkarılmasını sağlamaktadır. TBA ise bir boyut azaltma yöntemidir. Öznitelik çıkarım işlemi sonucu oluşan öznitelik matrisine öznitelik seçim işlemi uygulanmıştır. Öznitelik seçimi, öznitelik matrisinden sınıflandırma doğruluk oranını arttıran öznitelik sütunlarının seçilmesi ile oluşturulmaktadır. Öznitelik sütunu seçiminde farklı sütunların seçilme olasılığı vardır. Bütün olasılıkların tek tek denemesi zor bir işlemdir. Bu sebeple alternatifler arasından seçim yapılarak karar verildiği için bu işlem bir optimizasyon problemidir. Bu tezde optimizasyon algoritmalarından Diferansiyel Evrim Algoritması'nın öznitelik seçimi amacıyla (DEÖS) kullanılması incelenmiştir. Bu algoritma ile en yüksek sınıflandırma doğruluk oranına sahip alt kümeler seçilmiştir. Ayrıca öznitelik seçimi için istatistiksel bir test olan ki kare yöntemi de denenmiştir. Farklı veri kümeleri ile yapılan testlerde Bonn verisi için ADD'nin ürettiği öznitelik kümesi sınıflandırıcılarla test edildiğinde yüksek doğruluk oranları elde edilmektedir. Sağlıklı, hasta ve nöbet geçiren kişilerden alınmış 3'lü kümelemelerde literatürde görülen doğruluk oranlarını geçmektedir. Ayrıca nöbet geçiren ve geçirmeyen kişilerin verileri sınıflandırıldığında elde edilen doğruluk oranları da literatürdekileri geçmiştir. ADD ile elde edilen öznitelik matrisinden DEÖS kullanılarak en yüksek doğruluk oranını veren özniteliklerin kullanılması sınıflandırma doğruluk oranını arttırmaktadır. 1D-YİÖ yöntemi ikili ve üçlü kümelerde ADD'den bir nebze düşük sonuçlar vermesine rağmen kendisi gibi olan literatür çalışmasıyla yarışabilmektedir. Fakat beşli kümelemelere gelindiğinde ADD ile yarışmamaktadır.

# **EFFECTS OF FEATURE EXTRACTION AND DIFFERENTIAL EVOLUTION FEATURE SELECTION METHODS ON CLASSIFICATION SUCCESS OF EEG SIGNALS**

## **SUMMARY**

EEG signals are obtained by measuring the electrical activity in the brain. The analysis, examination and classification of these signals are used to diagnose disorders such as epilepsy and sleep disorder. It is sometimes difficult and complicated process to achieve a result by classifying the EEG signals and introducing logical approaches to these results. In this thesis, studies were carried out by feature extraction and selection methods based on pattern recognition to increase the classification accuracy of the EEG signal. In this thesis, the band pass filter was used to eliminate the noisy data in order to increase the classification accuracy. After the filtering process, data extraction methods were applied to the data. After the attribute extraction process, the attributes that increase the classification accuracy are selected by the feature selection process. The obtained attribute matrix was tested by Support Vector Machines (SVM) and Linear Discriminant Analysis (LDA). Classification accuracy were compared. The specified methods were applied on two different data. The specified methods were applied on two different set of data. The first one is the data collected by the University of Bonn for the detection of epilepsy. This data set has 100 single-channel samples and consists of 5 different classes. 2 of these classes are healthy, 2 patients and 1 patient are seizures. The other one is BCI competition data collected by Berlin Technical University for the study of muscle activities. 3a data set of the 3rd competition of BCI was used. This data set has 60 channels and consists of 4 different classes. These classes represent the data of the right hand, left hand, tongue and foot movements. The feature extraction process is to adapt the data to the classification process by using different filtering and statistical methods from the existing data set. During this process, the column structures are changing and a new set is created to represent the data. Common Spatial Patterns (CSP), 1D Local Binary Pattern (1D-LBP), Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) were used for feature extraction methods. CSP from feature extraction methods is a spatial filtering method commonly used to distinguish separate classes and to increase the spatial resolution of the signals obtained. CSP is used for two-class clusters, but in this thesis the extended state is used for multi-class clusters. 1D-LBP is a method

derived from image processing methods such as many of the signal processing methods. It treats the signal like an image and arranges the data according to whether the central signal is bigger than the neighboring signals. DWT allows the data to be separated by dividing the signal into different subbands and by filtering the data in the desired frequency range and using the coefficient information of the data in that range. PCA used for attribute extraction is a size reduction method. The feature selection process was applied to the feature matrix resulting from the attribute extraction process. The feature selection is created by selecting feature columns that increase the accuracy of the classification from the feature matrix. Columns of subset could be selected by different combination in feature selection. There is many possibility. It is difficult to try all possibilities individually. For this reason, this process is an optimization problem because it is decided by choosing from alternatives. In this thesis, the use of Differential Evolution Algorithm (DEFS) for optimization of optimization algorithms is investigated. The subscales with the highest classification accuracy were selected with this algorithm. In addition, chi-square method, which is a statistical test for attribute selection, has also been tried. For tests conducted with different data sets, high accuracy rates are obtained when tested with the attribute set classifiers produced by DWT for Bonn data. Healthy, patients and seizures in the 3-group clusters taken from people in the literature exceeds the accuracy rates. In addition, the accuracy obtained when the data of the people who had seizures and not passed were passed in the literature. The use of attributes that give the highest accuracy by using DEFS from the attribute matrix obtained by DWT increases the classification accuracy rate. Although 1D-LBP method has some slightly lower results than DWT in binary and triple clusters, it can compete with literature. However, it cannot compete with DWT when it comes to clusters of five.





## 1. GİRİŞ

Beyin çalışması esnasında düşük şiddette elektrik akım üretir. Beyin Bilgisayar Arayüzü (BBA) kişilerin beyinleri tarafından üretilen elektriksel akımın ölçülmesi ve bilgisayarlar yardımıyla yorumlanabilmesi işlemidir. BBA çeşitli sebeplerle beyin fonksiyonlarını kullanabilen fakat kas sistemini yarım ya da tamamıyla kullanamayan bireyler için düşünce gücüyle bir sistemi, bir bilgisayarı ya da bir protezi harekete geçirmek için kullanılan sistemlerdir. EEG (Elektroensefalogram) beyindeki elektriksel akımın kafaya bağlanan elektrotlar sayesinde beyin sinyallerinin bilgisayar ekranına aktarılması ile elde edilen bir BBA türüdür.

Bu çalışmanın amacı EEG verilerinin sınıflandırma doğruluk oranlarının artırılmasıdır. Bu amaç doğrultusunda farklı öznitelik çıkarım ve seçim yöntemlerinin sınıflandırma doğruluğuna katkısı incelenmiş ve birbirleriyle karşılaştırılmıştır. Bu tezde yapılan işlemler sırasıyla filtreleme, öznitelik çıkarımı, öznitelik seçimi ve sınıflandırma işlemidir. İlk olarak kullanılan verilere filtreleme işlemi uygulanmıştır. Veriler toplanma amaçlarına göre farklı frekans aralıklarından elde edilmiş olabilmektedir. Bu sebeple filtreleme işlemi verilere göre farklı frekans aralıklarında seçilmiştir. Filtreleme işleminin detayları Ön işleme ve Gürültü Temizleme kısmında anlatılmıştır. Filtrelenen veriye öznitelik çıkarım işlemi uygulanmıştır. Bu işlem için kullanılan yöntemler CSP, ADD, 1D- YİÖ ve TBA'dır. Bu metodların detayları öznitelik çıkarma ve boyut indirgeme kısımlarında anlatılmıştır. Öznitelik çıkarımı işleminden sonra öznitelik seçimi uygulanmıştır. Öznitelik seçimi alt küme seçimi olarak da bilinen bir yöntemdir. Öznitelik matrisinden sınıflandırma doğruluk oranını arttıran öznitelik sütunlarının seçilmesi işleminde farklı sütunların seçilme olasılığı vardır. Bunun için iki yöntem kullanılmıştır. İlki ki-kare yöntemidir. Diğeri ise Diferansiyel Evrim algoritmasıdır. Bu yöntemler ile en yüksek sınıflandırma doğruluk oranına sahip alt kümeler seçilmiştir. Sonuçta elde edilen öznitelik matrisleri sınıflandırıcılar ile test edilerek sınıflandırma doğruluk oranları bulunmuştur. Sınıflandırma işleminde 10-k çapraz doğrulama işlemi uygulanmıştır. Doğruluk oranlarının ortalaması ve standart sapması hesaplanarak sonuçlar karşılaştırılmıştır.

Bütün işlemler literatürde yaygın olarak kullanılan Bonn Üniversitesi'nin halka açık epilepsi verisi [1] ve BCI üçüncü yarışmasına ait 3a veri seti [2] için uygulanmıştır. Bu verilerin detaylı incelenmesi ileriki bölümlerde anlatılmıştır.

### 1.1. EEG Nedir?

EEG, BBA'nın kullanılabilmesi için geliştirilmiş bir yöntemdir. Beyindeki elektriksel aktivitenin varlığını ilk defa Psikiyatrist Hans Berger bulmuştur ve 1924 yılında alfa ve beta dalgalarını kaydetmeyi başarmıştır [3]. EEG ölçümü güvenli, pratik ve kolay bir işlemdir. Kafatasına yerleştirilen elektrotlar sayesinde beyindeki sinyal durumunun rahatça ölçülmesini sağlar. Şekil 1.1'de EEG ölçümü yapılan bir birey gösterilmektedir [4].

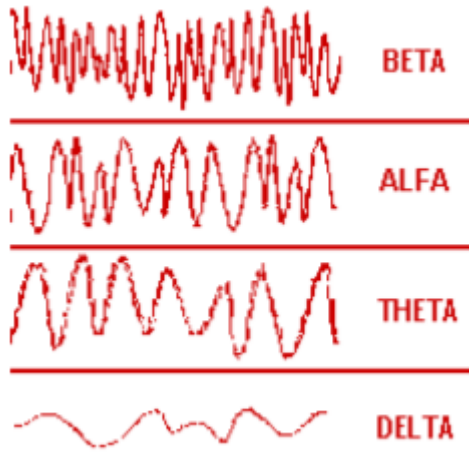


**Şekil 1.1:** EEG ölçümü örneği [4]

EEG verisinden faydalanma özellikle son zamanlarda kişinin beyin yapısının anlaşılması ve o kişide nörolojik bir bozukluk olup olmadığının takibi gibi konularda kullanılmaktadır. Çeşitli uyku bozuklukları, psikolojik rahatsızlıklar, epilepsi gibi hastalıkların tanı ve tedavisinde bu sinyallerin analizi büyük önem arz etmektedir. Kısmen ya da tamamıyla felçli ya da sadece beyin fonksiyonlarıyla hayatını idare eden bireylerin yaşamlarını daha rahat kontrol altına alabilmeleri için gerekli protezlerin ya da çeşitli tıbbi materyalin üretilmesi ve hasta ile uyum sağlayabilmesinde kullanılmaktadır. Ayrıca son yıllarda özellikle yaygınlaşan ve önemi anlaşılan pazarlama alanında yaygın olarak kullanılmaktadır. Bilinçaltına hitap etme çeşitli görsellerle kişilere istenildiği gibi kararlar aldırılmakta etkili olmaktadır. Bu konuda ilerleme kaydedebilmek için yapılan çalışmalar kişilerin karar verme mekanizmasını çözebilmek ve bir konuda fikir sahibi olduğu anda beynin ürettiği sinyalleri

sınıflandırabilmek için yapılan çalışmalardır. Son yıllarda oldukça önem kazanan *nöropazarlama* olarak bilinen kişilerin nasıl karar verebildiği, hangi aşamada ikna oldukları BBA ile tespit edilmiş ve ikna oldukları o andaki EEG verileri incelenmiştir. Bu sayede ikna olunan andaki durumlar pazarlama yapılmak istendiğinde en uygun konuma getirilerek kişinin yönlendirilmesi planlanmaktadır.

EEG sinyalleri belirli genlik ve frekansa sahip periyodik olmayan işaretlerdir. Bu sinyaller kişinin durumuna bağlı olarak çeşitlilik gösterebilmektedir. EEG sinyalleri frekanslarına göre çeşitli alt bantlarına ayrılabilir. Uyanık olup olmama, kişinin psikolojik durumu, düşündüğü eylem, sahip olduğu nörolojik hastalık EEG sinyallerinin çeşitlenmesine sebep olur çünkü, kişinin beynindeki elektriksel aktivitesi uyurken, mutlu iken, depresyonda iken, hayal kurarken farklılık göstermektedir. Bu çeşitlenme ile birlikte çeşitli frekans bantlarında dalgalar oluşur. Bu dalgalar delta, teta, alfa, beta ve gama'dır. Alfa Beta, Teta ve Delta dalgalara ait frekanslar Şekil 1.2'de gösterilmiştir [5].



**Şekil 1.2:** Alfa Beta, Teta, Delta beyin dalgalarına ait frekanslar [5].

### **1.1.1.Teta Dalgaları:**

Frekansları 3-8 Hz aralığında genlikleri 5-100 mikrovolttur [6]. Derin uyku, Rem ve hayal gücü ile ilişkilidir. Uyuyan çocuklarda ve 13 yaş ve altı çocuklarda görülür. Stres altındaki kişilerde sıklıkla görülür [7]. Yetişkinlerde sıkça görülmesi bir rahatsızlığın temeli sayılabilir. Teta dalgasının sağlıklı seviyede olması yaratıcılık ve düşünmeyi olumlu etkiler [8].

### **1.1.2. Delta Dalgaları:**

3 Hz den daha küçük frekansa sahiptir. Genlikleri 20-400 mikrovolttur. Çocuklarda ve yetişkinlerde derin uyku anında görülür [6]. Genel olarak delta dalgaları kalp atış hızının düzenlenmesi gibi bilinçsiz vücut aktiviteleri ile ilgilidir [7]. Yetişkin bir bireyde bu dalgaların sık sık görülme durumu beyinde ciddi sorunların var olduğunu gösterir [8] .

### **1.1.3. Alfa Dalgaları:**

8-13 Hz aralığında frekansa ve genlikleri 2-10 mikrovolta sahiptir [6]. Vücudun rahatlamış anında ya da uykudan önceki gözler kapalı andaki gevşeme durumlarda aktiftir [7]. Stressiz bireylerde görülme ihtimali yüksektir [8].

### **1.1.4. Beta Dalgaları:**

12-33 Hz arası frekanslara ve genlikleri 1-5 mikrovolta sahiptir [6]. Kişinin uyanık olduğu ve zihinsel aktiviteler yapabildiği zamanda oluşan dalgalar [7]. Günlük hayatta düşünmek konuşmak çalışmak gibi durumlar beta dalgasının olduğu durumlardır [8]. Düşünmek, konuşmak, çalışmak vb. Yoğun çalışma durumunda, stresli anlarda, dikkati toplamanın çok zor olduğu zamanlarda bu dalgaya rastlanmaktadır. [9].

### **1.1.5. Gama Dalgaları**

Gama dalgaları genelde klinik çalışmalar esnasında dikkate alınmaz [6]. Diğer alt bantlara göre yüksek frekanslı bantlardır. Frekansları 30-80 Hz arasında genlikleri 2 mikrovolttan küçüktür [9].

## **1.2. Kaynak Araştırması**

EEG sinyallerinin incelenmesi için yapılmış birçok çalışma mevcuttur. Bunlar içinde farklı veri setleri üzerinde çalışma yapılanlar da vardır. Bu bakımdan kaynak özeti kısmında Bonn Üniversitesi'nin epilepsi verisi, BCI yarışması verisi ve diğer veri setleri şeklinde ayırım yapılarak incelenmiştir.

Bonn Üniversitesi'nin verisi beş adet veri kümesini içerir [1]. Bunlar A, B, C, D ve E'dir. A ve B sağlıklı C, D ve E verileri ise hasta deneklere aittir. En çok kıyaslanan A sağlıklı ve gözü açık denek ile E hasta ve epilepsi nöbeti halindeki denektir.

BCI verisi üçüncü yarışmasına ait 3a veri kümesidir [2]. Veri kümesinde 3 adet denek vardır. Denekler sandalyeye oturtulup ekranda gösterilen şekillere göre sağ el, sol el, ayak ve dillerini hareket ettirdiklerini düşünmeleri istenmektedir. Verilerin detaylı açıklaması veri işleme kısmında anlatılmıştır.

### **1.2.1. Bonn Üniversitesinin Epilepsi Verisiyle Yapılmış Çalışmalar**

Z. Lasefr ve arkadaşlarının 2017'de yaptığı çalışmada literatürde de sıklıkla kullanılan A ve E deneklerini kıyaslamışlardır. A deneği sağlıklı veriyi, E deneği epilepsi nöbet halindeki veriyi tanımlamaktadır. Bunun için filtreledikleri sinyali Ayırık Dalgacık Dönüşümü (ADD) yöntemiyle alt bantlara ayırmışlardır. Ayırdıkları alt bantların katsayılarından her birine ait enerji değerini hesaplamışlardır. Sınıflandırma için Destek Vektör Makinesi (DVM) ve Yapay Sinir Ağları (YSA) yöntemi kullanılmıştır. Sonuçta DVM %96 ve YSA %98 doğruluk oranı elde etmişlerdir.

Md. Rashid ve arkadaşının 2017'de yaptığı çalışmada [10] ADD ile Daubechies-4 (db4) dalgacığı kullanılarak beşinci seviyeye kadar ayrılmış ve 10 adet istatistik özellik çıkarmışlardır. Bunlar ortalama, medyan, maksimum, minimum, minimum aralık, standart sapma, ortalama mutlak sapma, ortanca mutlak sapma, l2 norm ve maksimum normdur. Sınıflandırma işlemini sinir ağları yöntemi ile ikili ve üçlü kombinasyonlar şeklinde ölçmüşlerdir (AxB, BxE, CxE, DxE, AxCxE, BxCxE, AxDxE, BxDxE). En yüksek %100 en düşük %79,3 bulmuşlardır.

A. Ahmadi ve arkadaşlarının 2017'de yaptığı çalışmada [11] sinyal 17 alt parçaya ayırmışlardır. Her bir parçaya ADD uygulamışlardır. ADD için dalgacık ailesinden Daubechies2, Daubechies4, sym4, rbio2.2, db6, bior2.4 ve bior1.1 gibi dalgacık aileleri kullanılmış ve sonuçları kıyaslamışlardır. En yüksek sınıflandırma doğruluk oranı bior1.1 ile elde etmişlerdir. Sınıflandırma doğruluk oranları bir sınıfa karşı ikinci sınıf ve bir sınıfa karşı bütün sınıflar şeklinde ikili sınıflandırmaya dayanarak DVM ile 2, 5, 10 k-katlamalı çapraz doğrulama uygulayarak incelemişlerdir.

W. Ren ve arkadaşlarının 2016'da yaptığı çalışmada [12] farklı öznitelik çıkarma metotları ile öznitelikler çıkarmışlardır. Bu yöntemler Otoregresif model (AR), ADD, Dalgacık Paket Dönüşümü (DPD) ve Örnek Entropi (SampEn)'dir. Bu yöntemlerle elde edilen özniteliklere Boyut azaltma ve öznitelik seçme metotları uygulanarak alt küme oluşturulmuşlardır. Oluşan alt kümeler ile sınıflandırma yapmışlardır. Bu yöntemler Temel Bileşen Analizi(TBA), Minimum Gereksizlik Maksimum İlgi (mRMR), Fisher Skoru. Elde edilen alt kümeler karar ağacı ve DVM ile sınıflandırma yapmışlardır. Sonuçta diğer çalışmalara kıyasla kendi önerdikleri öznitelik çıkarma metotlarının daha iyi sonuçlar verdiğini kaydetmişlerdir.

S. Ramakrishnan ve arkadaşının 2018'de yaptığı bir çalışmada [13] iki adet veri seti kullanılmıştır. Biri Bonn Üniversitesi'nin verisi [1] diğeri Boston Çocuk Hastanesi'nin verisidir [14]. Bu iki veri içinde ADD ile 4. Seviyeye kadar dalgacık ayrımı yapılmış ve dalgacık katsayıları elde edilmiştir. Her bir banttan elde edilen katsayılar için istatistiksel özellikler hesaplanıp bulanık tabanlı öznitelik seçim metodu uygulanmıştır. Sınıflandırma için verinin %50'si eğitim %50'si test için rastgele seçilmiştir. Önerdikleri bulanık tabanlı öznitelik çıkarımı metodu ile %95 doğruluk elde etmişlerdir.

O. Salem ve arkadaşlarının 2014'de yaptığı çalışmada [15] dördüncü seviye ADD uygulanarak alt bant katsayıları elde edilmiştir. Elde edilen katsayılardan minimum, maksimum, standart sapma ve ortalama özellikleri çıkarılıp Karınca Kolonisi Algoritması ile sınıflandırılmıştır. Nöbet halindeki hasta ile geri kalan sınıfları karşılaştırarak sınıflandırma yapmışlardır ve hata oranını %9 olarak bulmuşlardır.

S. Bose ve arkadaşlarının 2017 de yaptığı bir çalışmada [16] ADD ile EEG verileri alt bantlara ayrılmış ve istatistiksel özellikler çıkarılmıştır. Çıkarılan bu özellikler Çok Katmanlı Algılayıcı ve Rastgele Orman sınıflandırıcılarını kullanarak A, D ve E kümelerinin doğruluk oranı %98 olarak bulunmuştur.

D. Chen ve arkadaşlarının 2017'de yaptığı çalışmada [17] ADD'nin farklı dalgacık aileleri kullanılmıştır. Bunlar bior, sym, coif, rbio, haar 'dır. En yüksek sınıflandırma doğruluk oranının sym dalgacık ailesine üye sym2 dalgacığı ile elde edildiği görülmüştür. Çalışmada epilepsi nöbetindeki hastanın tespiti %99,33 doğruluk oranıyla ölçülmüştür.

Y. Kaya ve arkadaşlarının 2014’de yaptığı çalışmada [18] görüntü işlemede kullanılan yöntemin bir boyutlu sinyale uygulanması ile oluşan 1D-YİÖ (Yerel İkili Örüntü) kullanılmıştır. A ve E kümelerinin birbirinden sınıflandırıcı ile ayrılması %99 doğruluk oranıyla elde edilmiştir.

U. R. Acharya ve arkadaşlarının 2012’de yaptığı çalışmada [19] öznitelik çıkarımı için DPD ve TBA, öznitelik seçimi için ANOVA testi kullanılmıştır. Sınıflandırma için karar ağacı, k-en yakın komşu ve Gauss karışımı modeli kullanılmıştır. En yüksek sonuç %99 ile Gauss karışımı modelinden elde edilmiştir.

N. Ahammad ve arkadaşlarının 2014’de yaptığı çalışmada [20] ADD ile 4. seviyede ayırım yapılmıştır. Bunun için Daubechies 2 dalgacığı kullanılmıştır. Ayrılan alt-bantlardan (D3, D4 ve A4) istatistiksel özellikler çıkarılmıştır. AxExE sınıfları kıyaslanmıştır. Sınıflandırma doğruluk oranları %84,2’ye ulaşmıştır.

K. Mahajan ve arkadaşlarının 2011 de yaptığı bir çalışmada [21] ADD kullanılarak EEG sinyalleri alt-bantlara bölünmüş ve istatistiksel özellikler çıkarılmıştır. Temel bileşen analizi (TBA) ve Bağımsız Bileşen Analizi yöntemleriyle öznitelik boyutu azaltılmıştır. Yapay sinir ağları ile (YSA) epilepsi rahatsızlığı olan kişilerin ve sağlıklı kişilerin verileri sınıflandırılmıştır. TBA ile %93,63 ve Bağımsız Bileşen Analizi ile %96,75 olarak bulunmuştur.

A. Subasi ve arkadaşını 2010’da yaptığı çalışmada [22] ADD ile 5. seviye Daubechies 4 dalgacığı kullanılarak ayırım yapılmıştır. Her bir alt banttan öznitelik olarak ortalama, standart sapma, güç ve mutlak ortalama oran olmak üzere 4’er adet çıkarılmıştır. TBA, BBA (Bağımsız bileşen analizi) ve LDA (Lineer Diskriminant Analiz) yöntemleri ile boyut azaltma yapılmıştır. Ayrı ayrı DVM uygulanmış ve doğruluk oranları kıyaslanmıştır. En yüksek sınıflandırma doğruluk oranları LDA ile %100, BBA ile %99,5 ve TBA ile en düşük doğruluk oranı %98,75 elde edilmiştir.

### **1.2.2. BCI Yarışma Verisiyle Yapılmış Çalışmalar**

M. Z. Baig ve arkadaşlarının 2017’de yaptığı çalışmada [23] BCI 3. yarışmasına ait 4a veri seti [24] kullanılmıştır. Bu veri seti her bir denek için 118 kanal ve 280 deneme içerir. Bu çalışmada sinyaller bant filtresi kullanılarak 0,05-200 Hz arası filtrelenmiştir. Daha sonra Butterworth bant geçiren filtre ile 8-30 Hz arası

filtrelenmiştir. Öznitelik çıkarımı için Ortak Uzamsal Örüntüler (CSP) metodu kullanılmıştır. Çıkarılan öznitelikler diferansiyel evrim, parçacık sürüsü optimizasyonu, karınca kolonisi algoritması ve yapay arı kolonisi gibi evrimsel tabanlı algoritmalar kullanılmış ve sonuçları kıyaslanmıştır. Çalışmalarında önerdikleri diferansiyel evrim tabanlı öznitelik seçim algoritmasının en optimal öznitelik alt kümesini seçtiğini ve dolayısıyla en yüksek sınıflandırma başarısına ulaştığı tespit edilmiştir.

Ö. Aydemir ve arkadaşlarının 2016' da yaptığı çalışmada [25] BCI 2. Yarışmasına ait 3. veri seti [26] kullanılmıştır. Fourier Dönüşümünün genelleştirilmiş bir hali olan Fast Walsh Hadamard Transform kullanılmıştır. Dönüşüm sonucunda seçilen kanallardan öznitelik çıkarılmış ve öznitelik matrisi oluşturulmuştur. K-NN (K- En Yakın Komşu), YSA ve DAA ile sınıflandırma sonuçları karşılaştırılmıştır. En yüksek doğruluk oranı YSA da en hızlı hesaplama ise DAA da tespit edilmiştir.

Y. Wang ve arkadaşlarının 2006'da yaptığı çalışmada [27] BCI 3.yarışmasına ait 4a veri seti kullanılmıştır. Sinyale 0,05-200 Hz arası bant filtresi uygulanmıştır. CSP uygulanarak aktif kanallar bulunmuştur. Bulunan aktif kanalların özellikleri çıkarılarak öznitelik veri alt kümesi oluşturulmuştur. Sınıflandırma için Fisher diskriminantı kullanılmıştır.

I. Koprinska'nın 2009 da yaptığı çalışmada [28] BCI üçüncü yarışmasına ait 3a isimli veri seti [2] kullanılmıştır. Bu çalışmada üç denek örneklerinin sahip olduğu 4 sınıf için öznitelik seçim metotları incelenmiş ve sınıflandırmadaki doğruluk oranları ölçülmüştür. Kullanılan bu öznitelik seçme metotları; Bilgi kazanımı (Information Gain-BK), Korelasyon tabanlı Öznitelik Seçimi (CFS), ReliefF, Tutarlılık (ConsiStency-Based Feature Selection) and 1R Ranking (1RR) 'dir. Bu öznitelik seçme metotlarının başarısı farklı sınıflandırma algoritmaları kullanılarak test edilmiştir. Tüm denekler düşünüldüğünde en düşük boyuta indirgeyen öznitelik seçim metodunun tutarlılık olduğu tespit edilmiştir. Bu metottan sonra en iyi boyut indirmenin CFS olduğu tespit edilmiştir.

### **1.2.3. Diğer Veri Setleri ile Yapılmış Çalışmalar**

I.A. Siradjuddin ve arkadaşlarının 2017'de yaptığı çalışmada [29] UCI veri seti [30] kullanılmıştır. ADD modeli ile alt bantlara ayrılan sinyalden yedi adet istatistiksel özellik çıkarılmıştır. Bunlar; maksimum, minimum, ortalama, standart sapma ve

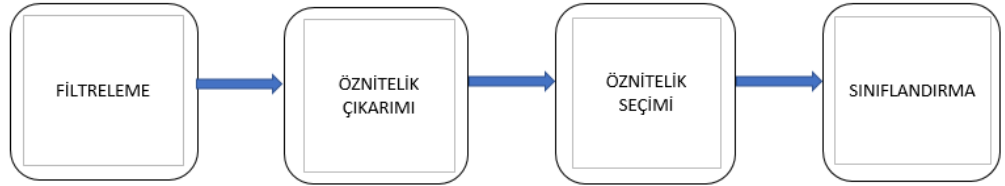


entropilerdir. Toplam öznitelik sayısı istatistiksel özelliğın kanal sayısına çarpımıyla elde edilmiştir. Sırasıyla verinin %20, %40 ve %60'lık kısmıyla eğitim yapılmış ve doğruluk oranları ölçülmüştür.

N.Ahammad ve arkadaşlarının 2014'de yaptığı çalışmada [20] kullanılan veri seti [31] için ADD ile 4. seviyede ayırım yapılmıştır. Bunun için Daubechies 2 dalgacığı kullanılmıştır. Ayrılan alt-bantlardan (D3, D4 ve A4) istatistiksel özellikler çıkarılmıştır. Sınıflandırma doğruluk oranı %98,5'ye ulaşmıştır.

## 2. VERİ İŞLEME VE ÖZNETELİK ÇIKARIMI

Bu tezde yapılan işlemlerin genel hatları Şekil 2.1’de gösterilmektedir. İlk olarak veri filtrelenmiştir. Filtrelenmiş veriye öznitelik çıkarım yöntemleri uygulanmıştır. Bu sayede veriyi temsil edecek öznitelik kümesi oluşturulmuştur. Daha sonra öznitelik kümesinden öznitelik seçimi yapılmıştır. Öznitelik seçimi sayesinde hangi özniteliklerin sınıflandırma doğruluk oranını arttırdığı tespit edilmiştir. Son aşamada ise öznitelik seçimi ile oluşmuş alt küme sınıflandırma algoritmaları ile test edilmiştir. Tezin bu aşamasında kullanılan verilerin tanımlanması ve filtreleme işlemleri anlatılmıştır.



Şekil 2.1: Tezde yapılan işlemler

### 2.1. Kullanılan Verilerin Tanımlanması

Bu tezde iki farklı veri seti üzerinde çalışılmıştır. İlk veri seti epilepsi rahatsızlığının tespiti için oldukça sık kullanılan Bonn Üniversitesinin verisidir [1]. İkinci veri seti ise BCI yarışması adı altında farklı yıllarda gerçekleştirilen yarışmada yayınlanan veri setlerinden biridir.

#### 2.1.1. Bonn Üniversitesi Verisi

Bu veri seti 5 kümeden oluşmaktadır (A, B, C, D, E). Her biri 23,6 saniyelik 100 örnekli tek kanallı kümelerdir. Örnekleme frekansı 173,61 Hz’dir. Her bir veri örneği 4096 veri noktası içermektedir. Bu veri setinde bozuk ve hatalı veriden

kaçınmak için denekler incelenmiş ve çeşitli kas hareketi ve göz kırpması gibi eylemlerin olduğu yerler kesilmiştir. Çünkü bu kas hareketleri beyinde ayrı bir elektriksel akım oluşmasına sebep olmaktadır. Bu durum esas incelenmek istenen sinyalin tespit edilememesine sebep olmaktadır. Ayrıca veri 0,5-85 Hz arası filtrelenmiştir. A ve B beş sağlıklı gönüllüden alınan verilerdir. A'daki gönüllülerin gözleri açık B'deki gönüllülerin gözleri kapalıdır. C, D, E ise beş epilepsi hastasının kayıtlarıdır. C sınıfına ait denekler EEG sinyali ölçümü sırasında kafatasına yerleştirilen elektrotların Epileptojenik bölgenin karşısına yerleştirilen elektrotla, nöbet halinde bulunmadıkları andan alınmıştır. D sınıfına ait denekler Epileptojenik bölgeye yerleştirilen elektrotlarla, nöbet halinde bulunmadıkları andan alınmıştır. E sınıfına ait örnekler ise, hastalar nöbet geçirirken alınmıştır. Verinin özeti Tablo 2.1'de gösterilmektedir [32].

**Tablo 2.1:** Bonn verisi özeti [32].

Denek	5 Sağlıklı Denek		5 Hasta Denek		
	Sınıf A	Sınıf B	Sınıf C	Sınıf D	Sınıf E
<b>Kişinin Durumu</b>	Uyanık ve gözler açık (Normal)	Uyanık ve gözler kapalı (Normal)	Hasta ve Nöbet Halinde değil	Hasta ve nöbet halinde değil	Hasta ve nöbet halinde
<b>Elektrot Tipi</b>	Yüzey	Yüzey	Kafatası içi	Kafatası içi	Kafatası içi
<b>Elektrot Yerleştirme</b>	Uluslararası 10-20 sistem	Uluslararası 10-20 sistem	Epileptojenik bölgenin karşısı	Epileptojenik bölge	Epileptojenik bölge
<b>Deneme Sayısı</b>	100	100	100	100	100
<b>Deneme Süresi</b>	23,6	23,6	23,6	23,6	23,6

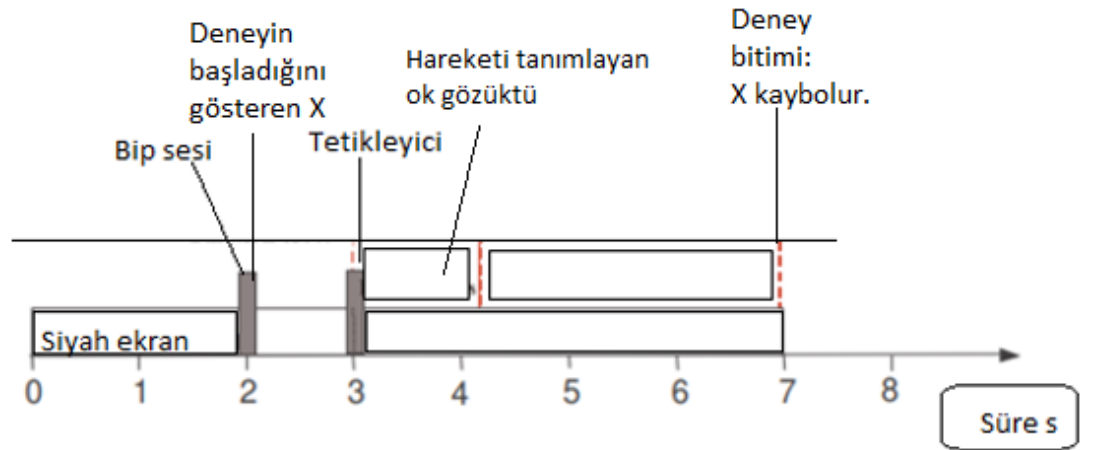
Tablo 2.1'de gösterildiği üzere temel olarak veri hasta ve sağlıklı kişiler olmak üzere ikiye ayrılmıştır. A'nın B'den tek farkının gözlerinin açık olmasıdır. Bu sebeple literatürde A ve B örnekleri için sınıflandırma doğruluk oranlarının karşılaştırılması nadiren görülmektedir. C, D ve E'nin hasta olduğu bilinmektedir. E diğer deneklerden önemli bir farka sahiptir. O da hasta kişiden nöbet geçirdiği anda alınan bir veri olmasıdır. Bu sebeple literatürde genelde E ile sağlıklı bireylerin (A, B) sınıflandırma doğruluk oranlarının kıyaslanması yapılmaktadır. Bu çalışmada bütün örnekler birbirleri ile ikili, üçlü ve dördü gruplar oluşturularak yeni kümeler elde edilmiştir. Elde

edilen kümelere sınıf etiketleri eklenmiştir. Etiketli kümelere sınıflandırma yöntemleri uygulanmıştır ve sınıflandırma doğruluk oranları karşılaştırılmıştır.

### **2.1.2. BCI Yarışma Verisi**

Bu çalışmada kullanılan veri setlerinden diğeri de BCI yarışmasının üçüncü yarışmasına ait 3a veri setidir [2]. Bu yarışmanın oluşturulma tarihi 2004'dür. Yarışma 8 veri setinden oluşmaktadır. Yarışmacıların çözüm bulması gereken konu ise eğitim ve test olarak ayrılmış verilerden sınıflandırma doğruluk oranı en yüksek sonucu bulmaktır.

3a veri setinin girdileri 64-kanallı bir EEG yükselteci ile yapılmıştır. Örnekleme frekansı 250 Hz'dir. 1-50 Hz arasında filtreleme işlemi uygulanmıştır. Deneyler şu şekilde gerçekleşmiştir; denekler kolçakları olan rahat bir sandalyeye oturtulmuş ve deneklerden ekranda gösterilen oklara göre sağ el, sol el, dil ve ayaklarını hareket ettirdiklerini düşünmeleri istenmiştir. Düşünülmesi istenen hareketlerin sırası rastgele olarak verilmiştir. Her bir hareket için 45'er deneme oluşturulmuş ve her bir denemede deneme başladıktan sonra 2 saniye sessizlik oluşmuştur. Sessizlik oluştuğu anda siyah ekran görülmüştür. İkinci saniyede ekranda '+' işareti ile bip sesi duyulmuştur. Üçüncü saniyede tetiklenme başlamıştır ve ekranda görülen bir saniyeliğine yukarı, aşağı, sağ ya da sola doğru bir ok gösterilmiştir. Bu anda '+' işareti kaybolana kadar denekten sağ el, sol el, dil ya da ayak hareketlerini ekranda gösterildiği sırayla hayal etmesi istenmiştir. Bu işlem yedinci saniyenin sonunda bitmiştir ve sekizinci saniyeye kadar deneğin dinlenmesi için vücudunu rahat bırakması istenmiştir. Deney bu şekilde toplam sekiz saniye sürmüştür. Deneyin gerçekleşme aşaması Şekil 2.2'de belirtilmiştir [2].



**Şekil 2.2:** Deneyin gerçekleştirme aşamaları [2].

Yarışmanın web sayfasından indirilen veri karmaşık görülmektedir. Bu sebeple ilk önce tetiklenme başlangıç ve bitiş örnekleri ele alınmıştır. İlk denek için indirilen dosyada trigger dosyası içinde tetiklenme başlama örneği verilmektedir. Bu zamanlar arasında 2560'lık bir fark olduğu görülmektedir. Tablo 2.2'de tetiklenme başlama örneğinin bir kısmı gösterilmektedir [2].

**Tablo 2.2:** Tetiklenme başlama örneği [2].

Tetikleyici
2425
4985
7545
10105
12666
15226

Tablo 2.2 bu hali ile deneyin istenilen kısmı ile ilgili bir anlam ifade etmemektedir. Mantıklı kısım Şekil 2.2’de gösterildiği gibi tetiklenme başlangıcından tetiklenme sonuna kadar olan kısımdır. Bu sebeple 3 ve 7 saniye arası dikkate alınmıştır. Aradaki 4 saniyede örnekleme oranı kadar örnek bulunmaktadır. Tablo 2.2’den elde edilen yeni tablo şu şekildedir; tetiklenme süresi eklenmesi ile (4 x250 Hz=1000 örnek) Tablo 2.3 elde edilmiştir [2].

**Tablo 2.3:** Tetikleyici işlem örneği [2].

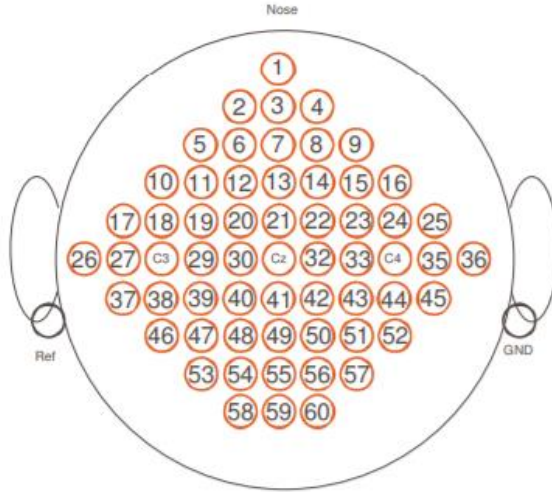
Tetiklenme Başlangıç	Deney Süresi	Tetiklenme Bitiş
2425	1000	3425
4985	1000	5985
7545	1000	8545
10105	1000	11105
12666	1000	13666
15226	1000	16226
17786	1000	18786
20346	1000	21346

Tablo 2.3’de yapılan işlem şu şekildedir. İlk olarak 8 saniyelik deneme süresinin asıl kısmı olan (tetiklenme başlangıcından deneğin rahatlamasına kadar geçen süre) 4 saniyelik kısmı büyük veriden çıkarılmıştır. Elde edilen verilerde anlamlı satırları alabilmek için 4 saniyelik kısım için örnek sayısı bulunmuştur. Büyük veri dosyasından tetiklenme bitiş ve başlangıç arası değerler alınmıştır. Veri dosyası bu şekilde düzenlenmiştir. Veri sayfasında bulunan doğru etiketli veriler ile elde edilen satırlar eşleştirilmiştir. Böylece denekler için aşağıdaki Tablo 2.4 elde edilmiştir [2]. Bu tabloya göre 1. deneğin 360 örneği bulunmaktadır. Kullanılan metotlar 1. deneğin verisinde denenmiştir.

**Tablo 2.4:** Denekler için eğitim ve test sayıları [2].

Denekler	Eğitim Sayısı	Test Sayısı
1. Denek	180	180
2. Denek	120	120
3. Denek	120	120

EEG elektrotlarının kafaya yerleştirilmesi Şekil 2.3'deki gibidir [2]. Şekilde gösterildiği üzere elektrotlar kafanın ön tarafından üçgen şeklinde genişleyerek yerleştirilmiş ve merkez hizasında daralarak enseye doğru ilerlemiştir. Yerleştirilen elektrotlar verideki kanalları temsil etmektedir. Bazı kanallardan elde edilen bilgiler kafaya yerleştirildikleri konumdan dolayı daha iyi bilgi vermektedir çünkü, bazı bölgeler çeşitli kas hareketlerinin beyindeki merkezine yakın olması sebebiyle kas hareketlerinden daha çok etkilenmektedir. Bu sebepten çok kanallı sistemlerde en doğru sinyallerin elde edilebilmesi için kanal seçimi önem kazanmıştır. Kanal seçimi ile ilgili çalışma öznitelik çıkarma kısmında incelenmiştir.

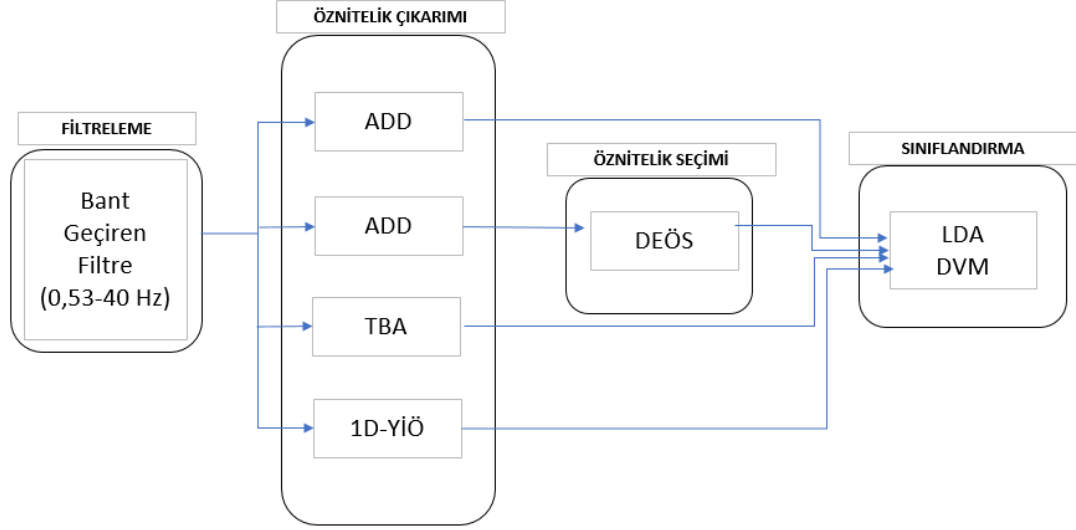


**Şekil 2.3:** Elektrotların kafadaki pozisyonları [2].

### 2.1.3. Ön işleme ve Gürültü Temizleme

Bu çalışmada Bonn verisi için uygulanacak işlemlerin genel diyagramı Şekil 2.4'de verilmiştir. Bu şekle göre ilk önce veri için filtreleme işlemi yapılmıştır. Daha sonra filtrelenmiş veriye ADD, TBA ve 1D-YİÖ uygulanarak öznitelik çıkarımı yapılmıştır.

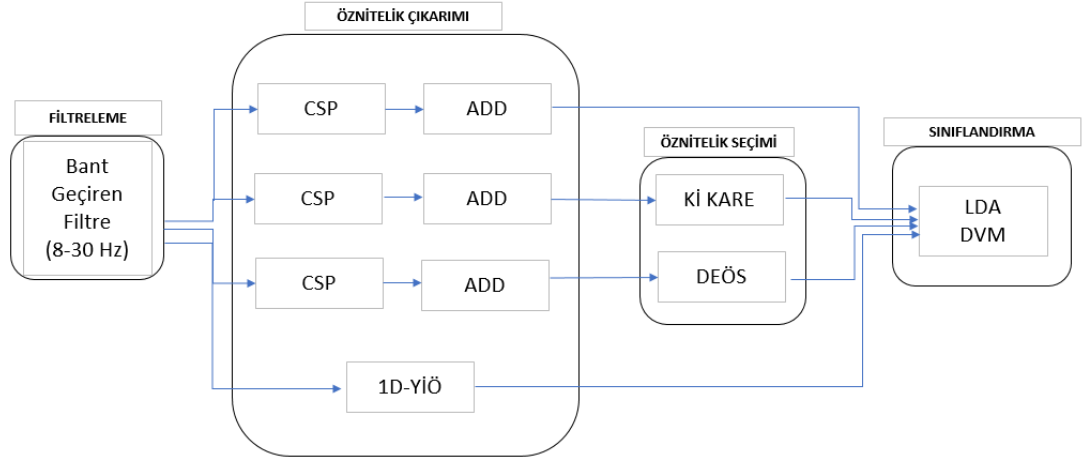
Öznitelik çıkarımı sonucu elde edilen öznitelik matrisine DEÖS ile öznitelik seçimi uygulanmıştır. Öznitelik seçimi sonucu sınıflandırma doğruluk oranını artıran öznitelik sütunları seçilmiştir. Seçilen sütunlar sınıflandırıcılar ile test edilmiş ve doğruluk oranları karşılaştırılmıştır. Bu işlemlerin teorik kısmı Öznitelik Çıkarma bölümünde, pratik kısmı ve ise Bulgular bölümünde incelenmiştir.



**Şekil 2.4:** Bonn verisi için uygulanan işlemler

BCI verisi için uygulanan yöntemler Bonn verisine uygulananlardan birkaç farka sahiptir. Bu işlemlerin gösterildiği diyagram Şekil 2.5'deki gibidir. BCI verisine ilk olarak filtreleme işlemi uygulanmıştır. Filtrelenmiş veriye ilk olarak CSP uygulanarak kanal seçimi yapılmıştır. Öznitelik çıkarımı için ADD ve 1D-YİÖ uygulanmıştır. Öznitelik çıkarımı sonucu oluşan öznitelik matrisine DEÖS ve ki kare ile öznitelik seçimi uygulanmıştır. Öznitelik seçimi sonucu oluşan öznitelik alt kümesi sınıflandırıcılar ile test edilmiştir ve doğruluk oranları karşılaştırılmıştır. Bu işlemlerin ayrıntılı anlatımı ve uygulanması ileriki bölümlerde anlatılmıştır.





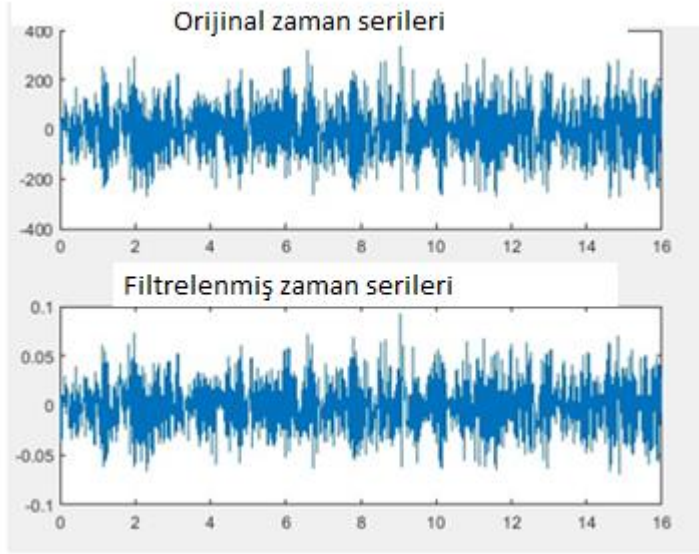
**Şekil 2.5:** BCI verisi için uygulanan işlemler

İlk olarak Bonn Üniversitesi verisine 0,53 ile 40 Hz arası filtreleme uygulanmıştır. Bu filtreleme işleminin benzer frekans aralıklarında literatürde başka çalışmalarda da uygulandığı Tablo 2.5’de görülmektedir.

**Tablo 2.5:** Literatürde benzer filtreleme örneklerinin görüldüğü çalışmalar.

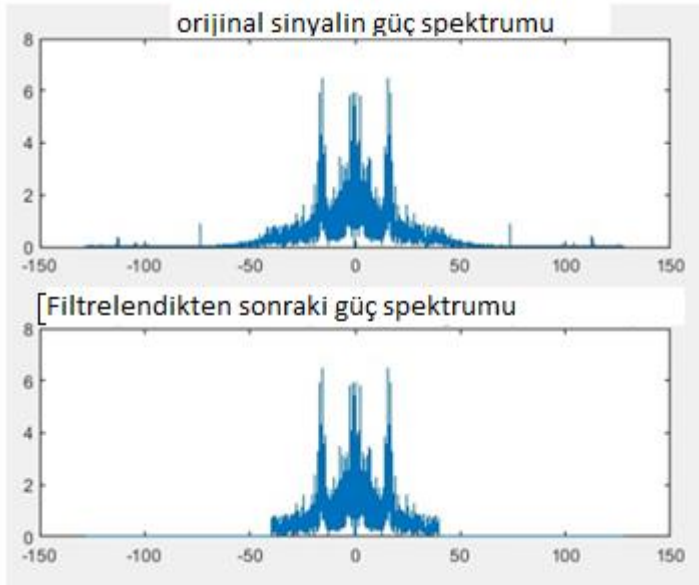
Çalışma	Filtreleme Aralığı
[13]	0-60 Hz
[17]	$\geq 0.5$ Hz.
[20]	0.53- 40 Hz

Tavsiye edilen ve iyi sonuç verdiği bilinen bu filtreleme aralığı kullanılarak veriye bant geçiren filtresi uygulanmıştır. Bant geçiren filtreler belirlenen üst ve alt limitler dışındaki diğer sinyalleri bastıran ve istenen aralıktaki sinyalleri ortaya çıkaran filtreleme türüdür. Filtre uygulanmadan önceki ve sonraki hali Şekil 2.6’daki gibi gösterilmektedir.



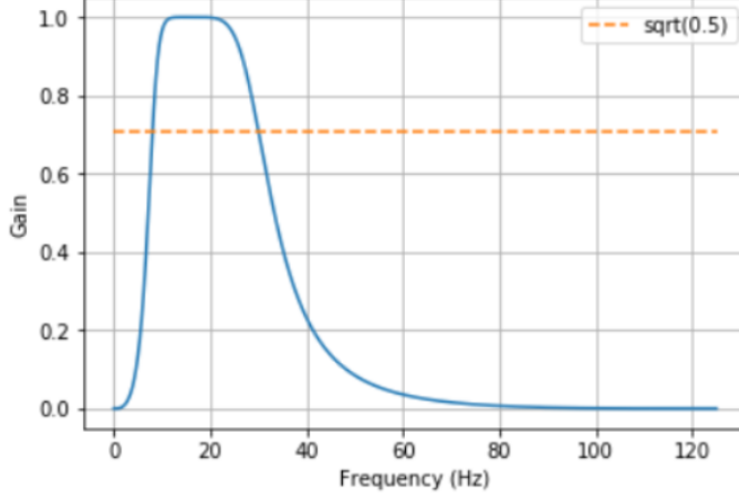
**Şekil 2.6:** Orijinal ve filtrelenmiş sinyalin frekans aralığı.

Filtrelenmeden önce ve filtrelenmeden sonra bant geçişliliğinin toplam gücü Şekil 2.7'deki gibi elde edilmiştir.



**Şekil 2.7:** Orijinal ve filtrelenmiş sinyalin bant geçişliliğinin toplam gücü.

BCI yarışma verisine 8-30 Hz arası bant geçiren filtre uygulanmıştır. Filtre uygulandıktan sonra belirlenen frekans aralığı dışındaki sinyaller bastırılmıştır. Sinyalin filtre uygulandıktan sonraki hali Şekil 2.8'deki gibi gösterilmektedir. Filtreleme işleminden sonra CSP işlemi uygulanmıştır.



**Şekil 2.8:** Bant geçiren filtre uygulandıktan sonra verinin frekans aralığı.

## 2.2. Öznitelik Çıkarma ve Boyut İndirgeme

Bu kısımda filtrelenmiş çok kanallı veriye kanal seçimi uygulanmıştır. Kanal seçimi uygulandıktan sonra iki veriye de öznitelik çıkarım yöntemleri uygulanarak öznitelik matrisi oluşturulmuştur.

### 2.2.1. Ortak Uzamsal Örüntüler (CSP)

Ortak uzamsal örüntü bulma yöntemi bir veriden çeşitli filtreler kullanarak öznitelik çıkarmak için kullanılmaktadır. Genelde iki sınıfı birbirinden ayırmak için kullanılmaktadır. Fakat çoklu sınıflar içinde genişletilmiş hali mevcuttur.

Bu tezde BCI yarışma verisi için uygulanan CSP şu şekilde gösterilmiştir.  $N$  kanal sayısını,  $M$  her bir deneme için örnek sayısını temsil etmiştir.  $S1$  ve  $S2$ ,  $N \times M$  boyutunda olmak üzere ayrı sınıfların verilerini temsil etmiştir. CSP işlemi için ilk önce her sınıfın uzamsal kovaryansı hesaplanmıştır. EEG'nin normalize edilmiş uzamsal kovaryansı birinci sınıf için Denklem 2.1 ve ikinci sınıf için Denklem 2.2 ile hesaplanmıştır.

$$X1 = \frac{(S1.S1^T)}{\text{trace}(S1.S1^T)} \quad (2.1)$$

$$X2 = \frac{(S2.S2^T)}{\text{trace}(S2.S2^T)} \quad (2.2)$$

$Trace()$  işleminin diagonal elemanların toplamını temsil etmiştir.  $S1^T$  ise  $S1$  matrisinin transpozunu temsil etmiştir. Uzamsal kovaryans matrisi her bir sınıf için her bir denemenin ortalaması alınıp hesaplanmıştır. BCI verisinde her bir deneğe 45 deneme yapıldığı için 45'er kez kovaryans matrisi bulunup, bulunan matrislerin ortalaması alınmıştır. Birinci ve ikinci sınıfların uzamsal kovaryansını gösteren matris Denklem 2.3'deki gibi gösterilmiştir.

$$\overline{X1}, \overline{X2} \quad (2.3)$$

Her bir sınıf için uzamsal kovaryans matrisi bulunduğundan sonra kompozit uzamsal kovaryans matrisi bulunur. Denklem 2.4'deki gibi hesaplanmıştır.

$$X = \overline{X1} + \overline{X2}$$

(2.4) Daha sonra kompozit matris denklemi hesaplanmıştır. Bunun için Denklem 2.5'deki gibi kompozit uzamsal kovaryans özdeğerleri ve özvektörlerine ayrılmıştır.

$$X = \overline{X1} + \overline{X2} = R_0 \lambda_0 R_0^T \quad (2.5)$$

Denklem 2.5'deki  $R_0$  özvektörlerin olduğu matrisi,  $\lambda_0$  ise özdeğerlerinin bulunduğu diagonal matrisi temsil etmiştir. Özdeğerler azalan sırada sıralanmıştır. Özdeğerler ve özvektörler ile beyazlatma dönüşümü matrisi Denklem 2.6'ya göre elde edilmiştir.

$$P = \sqrt{\lambda_0^{-1}} R_0^T \quad (2.6)$$

İki ayrı sınıfa ait kovaryans matrisleri Denklem 2.7'ye göre hesaplanmıştır. Bu işlem sonucunda ortak özvektörler elde edilmiştir.

$$U_1 = P \overline{X1} P^T \quad U_2 = P \overline{X2} P^T \quad (2.7)$$

$U_1$  ve  $U_2$  ortak özvektörler içeren matrisler olmuştur. İçerdikleri özdeğerlerin toplamı Denklem 2.8'de olduğu gibi 1 olmuştur.

$$U_1 = U \lambda_1 U^T, \quad U_2 = U \lambda_2 U^T, \quad \lambda_1 + \lambda_2 = I \quad (2.8)$$

$I$  birim matrisi temsil etmiştir. Projeksiyon matrisi de Denklem 2.9'a göre bulunmuştur.

$$W = U^T P \quad (2.9)$$

$W^{-1}$  matrisinin sütunları EEG'nin ortak uzamsal desenlerini oluşturmuştur. Öznitelik çıkarımı için  $W$ 'nin satırları kullanılarak varyans hesaplanmıştır. İlk ve son  $m$  matris

satırının varyans hesabı dikkate alınarak  $Z$  hesaplanmıştır. ( $Z_p: p = 1, 2 \dots, 2m$ ).  $var(2p)$  formülü varyansı temsil etmiştir. Öznitelik Denklem 2.10'daki gibi hesaplanır.

$$F_p = \log \left( \frac{var(2p)}{\sum_{i=1}^{2m} var(z_i)} \right) \quad (2.10)$$

Literatürde çoklu sınıf için birçok CSP yaklaşımı bulunmaktadır. Bunlar; bire karşı kalanı (BKK), ikili kombinasyonlar halinde olan yaklaşımlardır [33]. BKK yaklaşımı uygulandığında bir sınıfa karşı geriye kalan sınıflar bir bütün olarak ele alınır [34]. BCI veri seti için BKK uygulanıp 4 tane projeksiyon matrisi oluşturulmuştur;

- (1. sınıf) X (2,3,4). Sınıf
- (2. sınıf) X (1,3,4). Sınıf
- (3. sınıf) X (1,2,4). Sınıf
- (4. sınıf) X (1,2,3). Sınıf

İkinci yaklaşımda ise bütün sınıflardan ikililer oluşturup projeksiyon matrisleri hesaplanmıştır. BCI veri seti için uygulandığında projeksiyon matrisleri şu şekilde oluşmuştur;

- (1. sınıf) X (2. sınıf)
- (1. sınıf) X (3. sınıf)
- (1. sınıf) X (4. sınıf)
- (2. sınıf) X (3. sınıf)
- (2. sınıf) X (4. sınıf)
- (3. sınıf) X (4. sınıf)

Bu tezde BCI veri seti için ikinci yolu seçerek 4 sınıflı bir veri için 6 adet projeksiyon matrisi oluşturulmuştur. Elde edilen projeksiyon matrislerinden aktif kanallar belirlenmiştir. Aktif kanal seçiminde Denklem 2.11 uygulanmıştır. İlk sınıfa ait uzamsal desen  $S_1$ , ikinci sınıfa ait uzamsal desen  $S_2$  olduğunda;  $find()$  metodu ile  $W^{-1}$  'nin sütunlarında mutlak değeri (genliği) en yüksek olan elemanın kaçınıcı satır olduğu hesaplanmıştır. Bu hesapla hangi kanalın aktif olduğu ortaya çıkmıştır.

$$CH^1 = find(|S_1| == Max(|S_1|)) \quad (2.11)$$

Elimizde altı adet projeksiyon matrisi olduğu için her bir matrisin ilk 10 sütununun en yüksek genlikteki elemanının satır sayısı tespit edilmiştir. Ortak çıkan kanallara öncelik verilerek kanal seçimi yapılmıştır. Toplamda 60 kanallı veri 12 kanala düşürülmüştür. Bu işlem sonucunda seçilmiş kanallar şunlardır; 55, 28, 43, 44, 54, 41, 27, 60, 58, 23, 4, 37. Seçilen kanallara ayırık dalgacık dönüşümü uygulanmıştır.

### 2.2.2. Ayırık Dalgacık Dönüşümü (ADD)

Sinyalden daha fazla bilgi çıkarabilmek için zaman-frekans araçlarına başvurulmuştur. Dalgacık dönüşümü bir işaretin zaman-frekans analizi için kullanılan bir dönüşüm türüdür. Fourier yöntemleri durağan olmayan sinyaller için uygun olmadığından dalgacık dönüşümü metotları alternatif olarak kullanılmaya başlanmıştır [35]. Ayırık dalgacık dönüşümü sinyale uygulanan tekrarlayan analiz filtrelerinin uygulamasıyla sinyali temsil eden dalgacık katsayıları oluşturulmasıdır. Sinyale uygulanan filtreler yüksek geçiren ve alçak geçiren filtrelerdir. Dalgacıkları ayırmak için Daubechies dalgacık ailesinden db4 dalgacığı kullanılmıştır. Dalgacıklara ayırmak için db4'ten başka kullanılacak Sym, Coif gibi dalgacık aileleri de mevcuttur. Yapılan çalışmalar sonucunda Daubechies dalgacıklarının daha iyi sonuç verdiği ve db4'ün db2'den biraz daha iyi olduğu görülmüştür [36]. Bu sebeple bu tezde BCI yarışma verisi için db4 kullanılmıştır. Dalgacık dönüşümü matematiksel olarak Denklem 2.12'deki gibi ifade edilmiştir [36].

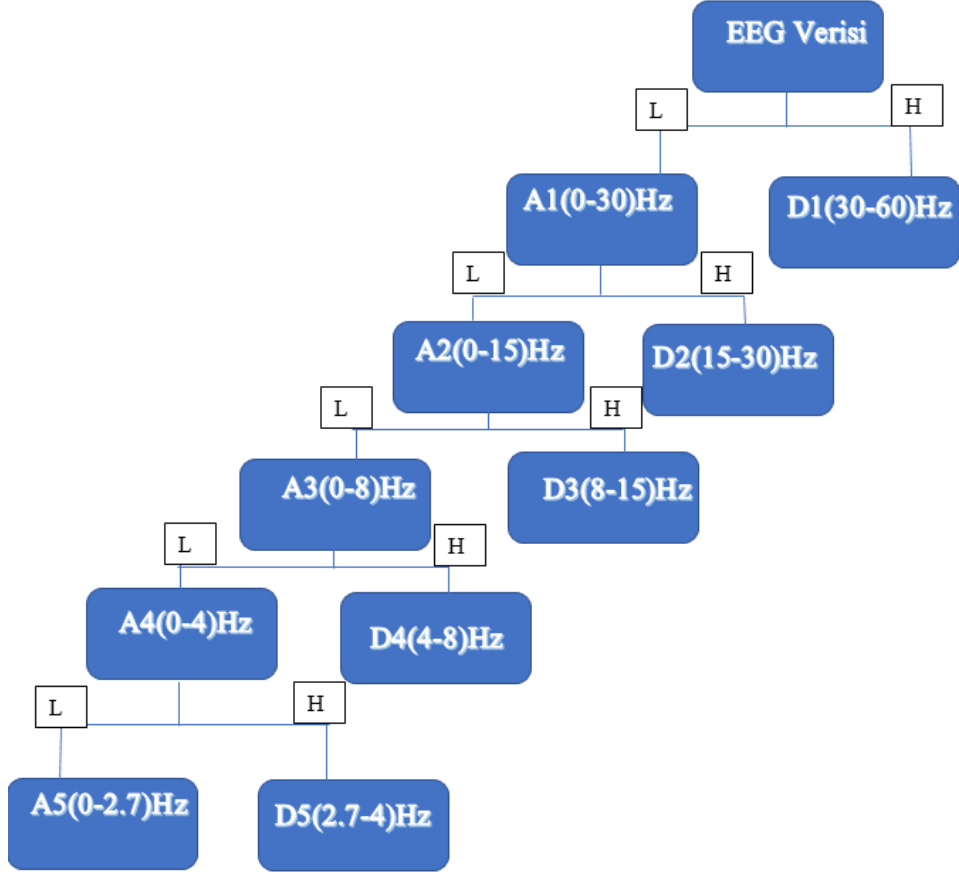
$$DWT(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi\left(\frac{t-b}{a}\right) dt, a = 2^j, b = k2^j = k \cdot a, j, k \in \mathbb{Z}^2 \quad (2.12)$$

$a$  değeri ölçekleme parametresini;  $b$  dönüşüm parametresini;  $\psi$  ana dalgacığı;  $DWT(a, b)$  işaretin sürekli dalgacık dönüşümünü belirtmiştir. Yüksek ve alçak geçiren filtreler de Denklem 2.13'deki gibi ifade edilmiştir.

$$y_{yüksek}[k] = \sum_n x[n] \cdot g[2k - n]$$

$$y_{alçak}[k] = \sum_n x[n] \cdot h[2k - n] \quad (2.13)$$

$g[n]$  yüksek geçiren  $h[n]$  alçak geçiren filtreyi ifade etmektedir. ADD beşinci dereceden dalgacık ayrışması Şekil 2.9'da gösterildiği gibi yüksek ve alçak geçiren filtre uygulanarak elde edilmiştir.



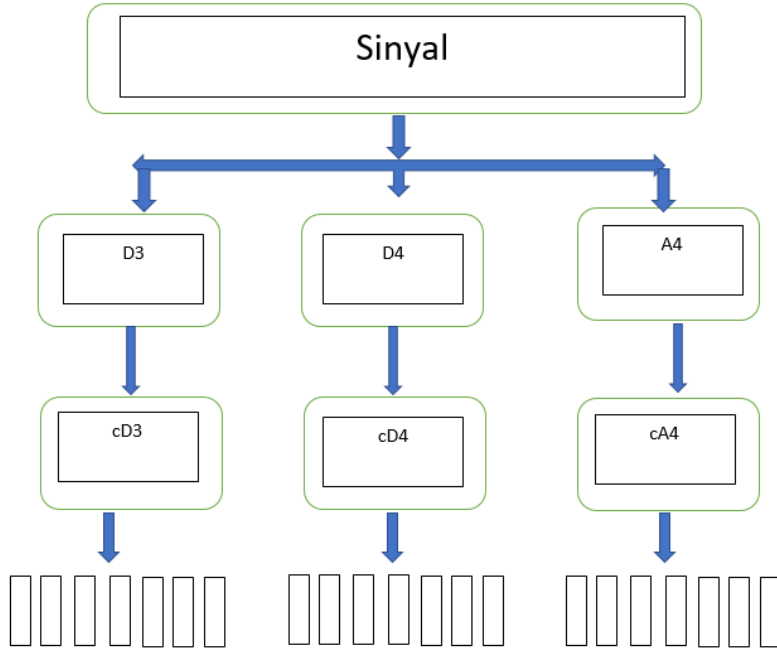
**Şekil 2.9:** EEG 5. Seviye Dalgacık Ayrışması.

Büyük veri ilk olarak yüksek ve alçak geçiren filtreler uygulanarak ikiye ayrılmıştır. Bir sinyale ait alçak geçiren frekans değerleri o sinyalin önemli özelliklerini taşımaktadır. Yüksek geçiren kısmı ise gürültülü veriyi taşımaktadır. Alçak frekans geçiren kısma tekrar filtre uygulanarak yeni seviyede de alçak ve yüksek frekans değerli sinyal oluşturulmuştur. Beşinci seviyeye kadar inildiğinde elde edilen katsayılar Şekil 2.9'daki gibi; D1, D2, D3, D4, D5 ve A5 olmuştur. Bu frekansların temsil ettiği dalgalar yaklaşık olarak Tablo 2.6'daki aralıklara denk düşmektedir.

**Tablo 2.6:** Dalgacık katsayılarının temsil ettiği dalgalar.

Dalga Adı	Dalgacık Katsayısı Adı
Delta	D5
Teta	D4
Alfa	D3
Beta	D2

ADD ilk olarak Bonn Üniversitesi verisi için daubechies 2 ile dördüncü seviyeye kadar ayrılmıştır. Dördüncü seviyede Şekil 2.6'da da gösterildiği üzere elde edilen katsayılar A4, D1, D2, D3 ve D4'tür. Sınıflandırmada kullanılan dalgacıklar D3, D4 ve A4'tir. Katsayıların temsil ettiği dalgalar Tablo 2.6'da gösterilmiştir. Literatürde yapılan çalışmalarda da alfa, delta ve teta dalgalarının seçildiği bilinmektedir. Seçilen üç alt banttan istatistiksel özellikler çıkarılmıştır. Bunlar her bir banttan elde edilen; maksimum, minimum, standart sapma, basıklık, çarpıklık, ortalama ve enerjidir. Böylece bir örnek için 8 (istatistiksel özellik) x 3 (alt bant) =24 öznitelik çıkarılmıştır. 100 örnek için verinin boyutu 100x24 olmuştur. Öznitelik kümesinin boyutu sınıf etiketleri eklenmesi ile bir sınıf için 100x25 olmuştur. ADD'den öznitelik çıkarma işlemi Şekil 2.10'da gösterilmiştir.



**Şekil 2.10:** ADD ile öznitelik çıkarımı

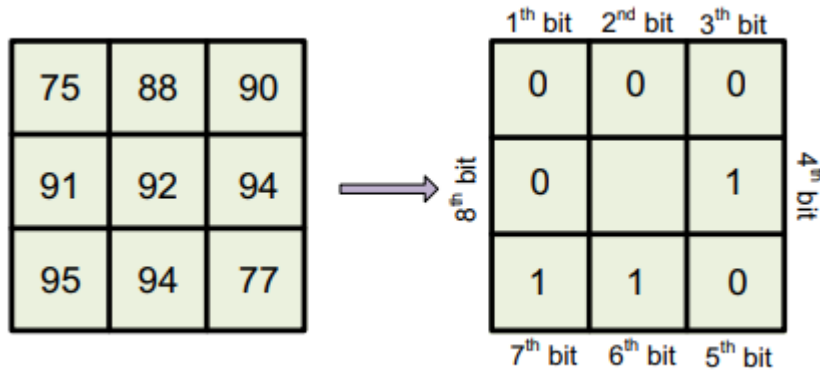
BCI yarışma verisi için daubechies 4 ile beşinci dereceden ayrıştırma uygulanmıştır. Elde edilen alt bantlar A5, D1, D2, D3, D4 ve D5 olmuştur. Seçilen alt bantlar A5, D5, D4 ve D3'tür. BCI yarışma verisi için ortalama, minimum, maksimum, enerji, eğrilik (skewness), basıklık (kurtosis), standart sapma, mutlak toplam ve varyans olmak üzere toplam 9 özellik çıkarılmıştır. 1 alt-banttan 9 özellik çıkarıldığında 4 alt-banttan  $9 \times 4 = 36$  özellik çıkarılmıştır. Bir kanaldan 36 öznitelik



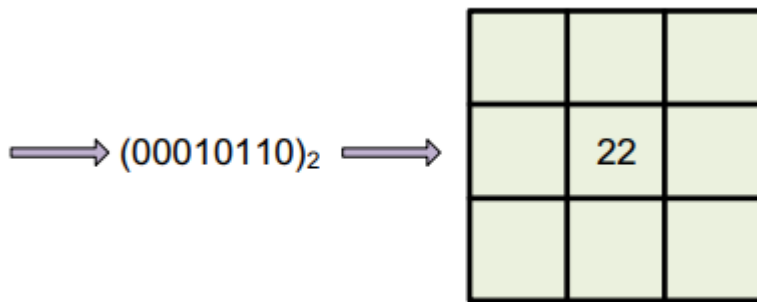
çıkarıldığında CSP ile seçilen 12 kanaldan  $12 \times 36 = 432$  öznitelik çıkarılmıştır. Bir sınıf için etiket sütunu da eklendiğinde  $180 \times 433$  boyutunda öznitelik kümesi elde edilmiştir.

### 2.2.3. Yerel İkili Örüntü (YİÖ)

Bu yöntem bu tezde ADD yöntemine alternatif olarak öznitelik çıkarımı için kullanılmıştır. YİÖ Ojala ve arkadaşları [37] tarafından bulunan yöntemdir. 1D-YİÖ yerel komşularda genel bir doku tanımından türetilen gri ölçekli bir doku ölçüsüdür [38]. 2D-YİÖ'da her bir piksel ve onun 8 komşusunun kıyaslanması ile oluşmaktadır. Ortadaki piksel değeri 8 komşusuyla kıyaslanmıştır. Denklem 2.14'deki gibi komşular büyükse bir değilse sıfır yazılarak iki tabanında bir sayı elde edilmiştir. 2D-YİÖ'nün görüntü üzerinde hesaplanması Şekil 2.11 [18] ve Şekil 2.12 [18]'de gösterilmiştir.



Şekil 2.11: Merkez pikselin komşu piksellerle değerinin kıyaslanması [18].



Şekil 2.12: Elde edilen 8 bitlik sayının 10'luk tabana çevrilmesi [18].

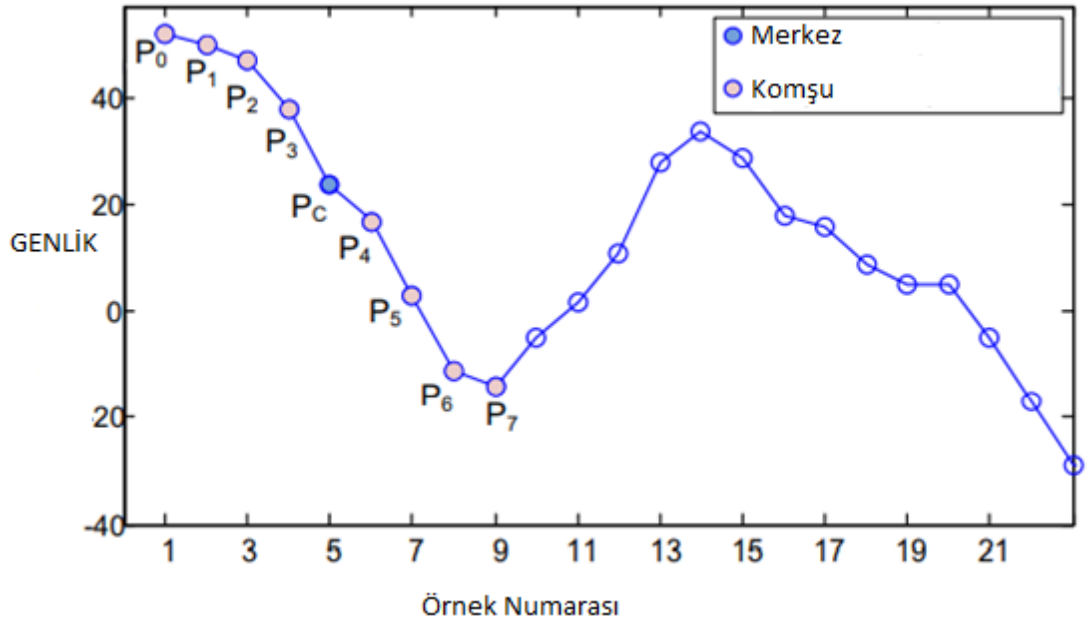
Denklem 2.14'deki  $G(x_i)$  merkezi pikselin  $i$ . komşusu ( $i = 1, 2, \dots, 8$ ),  $G(x)$  ise merkezi pikseli temsil etmiştir.

$$t = G(x_i) - G(x) \quad (2.14)$$

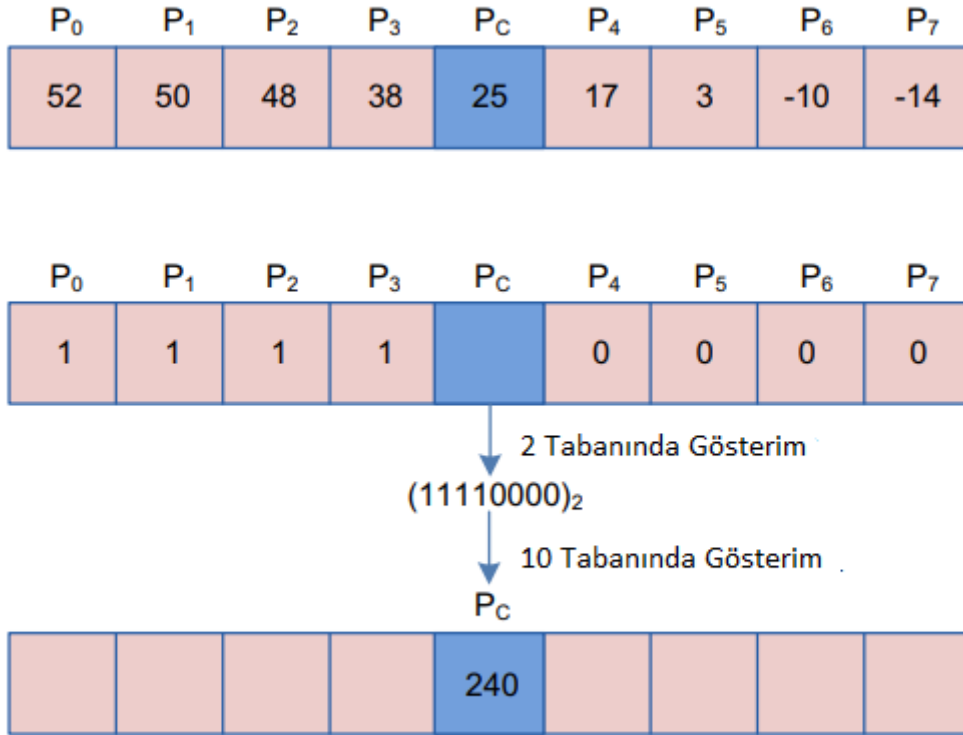
Sekiz bitlik sayı oluşturmak için Denklem 2.15 kullanılmıştır.

$$YİÖ(x) = \sum_{i=0}^p s(t) \cdot 2^i \quad , s(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (2.15)$$

2D-YİÖ'den türetilen 1D-YİÖ ilk olarak Chatlani ve arkadaşları [39] tarafından kullanılmıştır. 2D'de olduğu gibi hesaplama yapılmıştır. Merkezi değerle komşu değerlerin kıyaslanması sinyal uzunluğunca devam etmiştir. 1D-YİÖ'nün komşu sayısı 8 olarak belirlenip merkez sinyalinin ve komşu sinyallerin belirlenmesi Şekil 2.13 [18] ve 2.14'deki [18] gibi gösterilmiştir.



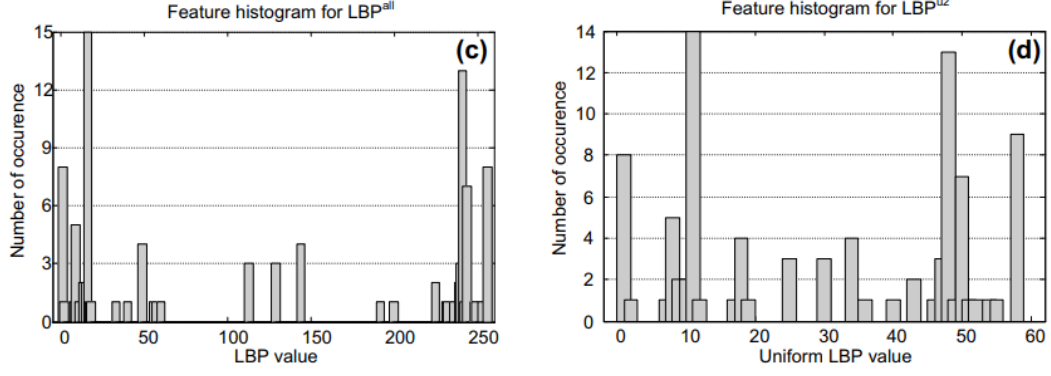
Şekil 2.13: Merkez sinyal ve kaç komşu sinyalin olacağına karar verme [18].



**Şekil 2.14:** 1d sinyale YİÖ uygulanması [18].

Bir kanal için 256 özellik elde edilmiştir. 256 özelliğin boyutunu düşürebilmek için 8 bitlik sayıların düzenli olup olmadığına bakılmıştır. Düzenli ve düzensiz sayılar şu şekilde hesaplanmıştır. Değişim sayısı 2 ve daha az ise düzenli fazla ise düzensizdir. Değişim sıfırdan bire birden sıfıra geçişleri ifade eder. Örneğin; 11100001 sayısı için 2 değişim olmuştur. Buna karşılık 11110101'da 4 değişim olmuştur. Bu durumda 11110101 sayısı düzensiz, 11100001 düzenlidir. Bu sayede 58 adet düzenli sayı olduğu bulunmuş ve her biri 58 aralığına yerleştirilmiştir. Düzensiz olup geriye kalan sayılar ise 59. sütuna yerleştirilmiştir. Daha önce 256 olan veri indeksi 59'a düşürülmüştür. Bu işlem Şekil 2.15'de gösterilmiştir [18].

Bonn verisi tek kanallı olduğundan bir örnek için 59 sütun elde edilmiştir. 100 örnek için 100x59 boyutunda öznitelik matrisi oluşmuştur. Sınıf etiketi eklendiğinde bir sınıf için 100x60 boyutunda öznitelik kümesi oluşmuştur.



**Şekil 2.15:** 256 aralığından 59 aralığına düşürülmesi [18].

BCI Yarışma verisi için 1D-YİÖ kullanılmıştır. Her bir kanaldan 256 özellik olmak üzere 60 kanal için  $256 \times 60 = 15360$  özellik çıkarılmıştır. Sınıf etiketi eklendiğinde  $180 \times 15361$  boyutunda öznitelik kümesi oluşmuştur.

#### 2.2.4. Temel Bileşen Analizi (TBA)

TBA çıktı bilgisini kullanmayan denetimsiz bir boyut indirgeme metodudur. Büyük boyutlu verilerin daha düşük boyutta özelliklerini kaybetmeden temsil edilebilmesini sağlar. 1901 yılında Pearson tarafından bulunmuştur [40]. TBA uygulama kolaylığından en çok kullanılan ve tercih edilen boyut indirgeme yöntemidir.

TBA kısaca  $x_1, x_2, \dots, x_n$  boyutlu bir verinin  $x_0$  vektörüyle ifade edilmeye çalışılmasına denir. Bu  $x_0$  vektörü ile yine aynı veriye ait  $x_k$  vektörünün farklarının mutlak değerinin karesi minimum olmalıdır. Hata değeri  $H_{error}$  ile ifade edildiğinde;

$$H_{error}(x_0) = \sum_{k=1}^n \|x_0 - x_k\|^2 \quad (2.16)$$

$H_{error}$  Denklem 2.16'daki gibi hesaplanmıştır. Bu işlem için  $x_0$  n adet vektörün ortalaması şeklinde hesaplanmıştır. Veriler örneğin ortalamasının etrafında olmak şartıyla bir doğruya yansıtılıp ve  $e$  yansıtılacak doğru ile aynı doğrultuda bir birim vektörü temsil ettiğinde ve  $a$  bir sayısal değer olduğunda doğru denklemi Denklem 2.17'deki gibi ifade edilir.

$$x = ort + a.e \quad (2.17)$$

Denklem 2.17 kısmı türev alınıp sıfıra eşitlendiğinde Denklem 2.18 elde edilmiştir.

$$a_k = e^T(x_k - ort) \quad (2.18)$$

Denklem 2.18 sonunda başka en iyi  $e$  değeri bulabilmek için saçılma matrisi kullanılmıştır. Saçılma matrisi Denklem 2.19'daki gibi hesaplanmıştır.

$$S = \sum_{k=1}^n (x_k - ort)(x_k - ort)^T \quad (2.19)$$

$a_k$  değeri Denklem 2.19'da yerine konulduğunda Denklem 2.20 elde edilmiştir.

$$H_0(e) = \sum_{k=1}^n ||ort - x_k||^2 - e^T S e \quad (2.20)$$

Lagrange çarpanları kullanılarak Denklem 2.21 elde edilmiştir.

$$u = e^T S e - (e^T e - 1) \lambda \quad (2.21)$$

Denklem 2.21'in  $e$  ye göre türevi alındığında Denklem 2.22 elde edilmiştir.

$$S e = \lambda e \quad (2.22)$$

Verinin tek bir boyuta yansımısını bulabilmek için en büyük özdeğere denk gelen özvektör seçilmiştir. Eldeki bu bir boyuta düşürme işlemi d-boyuta düşürmeye de dönüştürülebilmektedir. Bunun için formülde  $n$  yerine  $d$  konularak işlemler yapılabilmektedir.

### 2.2.5. Ki Kare Testi ile öznitelik seçimi

Ki Kare ( $\chi^2$ ) istatistiksel bir testtir. Gözlemlenen veriler ve seçilen verilerin birbiriyle uyumlu olup olmadığını inceleyen ve çıkan sonuca göre hipotezlerden birini kabul diğerini reddeden bir yöntemdir. Bu test için ilk önce iki adet hipotez gereklidir. Hipotezler  $H_0$  ve  $H_1$ 'dir.  $H_0$  hipotezi (null) sıfır hipotezdir. Doğru olduğu varsayılarak işleme başlanmıştır.  $H_1$  ise  $H_0$  'nin aksini savunan görüşü kabul etmektedir. Doğruluğu ispat edilmeye çalışılmıştır. İspat edilemezse  $H_1$ 'in yanlış  $H_0$ 'in doğru olduğu sonucuna ulaşılmaktadır. Ki kare testine başlanırken bir yanılma payı belirlenmektedir

(a). Sınıf sayısı  $c$  olarak kabul edilmektedir. Sınıftaki eleman sayısının belirlenmesi gerekmektedir. Denklem 2.23'deki gibi hesaplanmaktadır.

$$\Delta x = \frac{x_{mak} - x_{min}}{c} \quad (2.23)$$

$x_{mak}$ , örnekteki en büyük elemanı  $x_{min}$ , örnekteki en küçük elemanı ifade etmektedir. Bir sınıfın alt ve üst sınırı Denklem 2.23'den faydalanılarak şu şekilde hesaplanmaktadır.  $j = 1, 2, \dots, c$  kadar sınıf olduğu varsayılmaktadır.  $j$  ninci sınıfın alt sınırı ve üst sınırı Denklem 2.24 ve 2.25'deki gibi hesaplanmaktadır.

$$L_j = x_1 + (j - 1)\Delta x \quad (2.24)$$

$$U_j = x_1 + (j - 1)\Delta x \quad (2.25)$$

Her bir sınıf için gözlenen örnek değerleri ( $O_j$ ) Denklem 2.26'da gösterildiği şekilde ifade edilmektedir.

$$O_j = \sum_{i=1}^n \begin{cases} 1, & L_j \leq x_i < U_j \\ 0, & \text{diğerleri} \end{cases} \quad (2.26)$$

Her bir sınıf için beklenen örnek değerleri ( $E_j$ ) ile ifade edilmektedir. Bu durumda ki-kare değeri Denklem 2.27'deki gibi bulunmaktadır.

$$\chi_h^2 = \sum_{j=1}^c \frac{(O_j - E_j)^2}{E_j} \quad (2.27)$$

Hata payına göre  $\chi_{v,1-a}^2$  değeri ki-kare tablosundan bulunmaktadır.  $v$  değeri serbestlik katsayısını ifade etmektedir.  $\chi_h^2$  değeri  $\chi_{v,1-a}^2$ 'nın tablodan elde edilen değerden küçük ise  $H_0$  hipotezi red,  $H_1$  hipotezi kabul edilmektedir.

### **3. EEG VERİSİNİN SINIFLANDIRILMASI**

EEG verisinin sınıflandırılması için en çok tercih edilen sınıflandırma yöntemleri DVM ve LDA'dır. Bu tezde bu iki sınıflandırma metodu kullanılmıştır.

#### **3.1. Gözetimli Öğretim Yöntemi**

Etiketlenmiş veri üzerinden sistemin girdileri ve çıktıları yani etiketleri arasında mantıklı fonksiyonlar üreterek ilişki kuran yöntemlerdir. Bu tezde sınıflandırma işlemi için gözetimli öğretim metodlarından ikisi olan DVM ve LDA kullanılacaktır.

##### **3.1.1. Destek Vektör Makineleri (DVM)**

DVM büyük boyutlu sınıflandırma problemlerinde kullanılan önemli bir öğrenme algoritmalarından biridir. DVM uygulandığı veri setinin karmaşıklığını göz önünde bulundurarak modeli öğrenir böylece sınıflama ve regresyon yapar.

DVM'nin tarihi örüntü tanıma problemlerine 1936 yılında çözüm arayan R.A.Fisher'a dayanmaktadır. 1950 de ise Aronszajn DVM'nin parçası olan çekirdeği ortaya çıkarmıştır [41]. 1963'e gelindiğinde Vapnik ve Lerner DVM'nin temeli sayılabilecek Genelleştirilmiş Portre algoritmasını üretmişlerdir [42]. Daha sonraki zamanlarda istatistiksel öğrenme teorisinin temel adımları oluşturulmaya başlanmıştır. 1992'de Bernhard E. Boser, Isabelle M. Guyon ve Vladimir N. Vapnik maksimum marjlı hiperplanlara çekirdek kuralı uygulayarak lineer olmayan bir sınıflandırıcı ürettirler [43]. 1995 'de Corinna Cortes ve Vapnik tarafından yumuşak marj sınıflandırıcı geliştirilmiştir [44].

### 3.1.1.1. Doğrusal destek vektör makineleri

DVM ile sınıflandırmada amaç iki sınıfa ait örneklerin, eğitim verisi ile elde edilen bir karar fonksiyonu yardımıyla birbirinden ayrılmasıdır. Söz konusu karar fonksiyonu kullanılarak eğitim verisini en uygun şekilde ayırabilecek hiper-düzlem bulunmaktadır. Bahsedilen iki sınıfı ayırabilecek birçok hiper düzlem çizilebilir fakat DVM'nin amacı kendisine en yakın noktaların arasındaki uzaklığı maksimuma çıkaran hiper düzlemi bulmaktır. Bu sebepten noktalar arasında en uygun ayırımı yapan düzleme optimum hiper düzlem ve sınır genişliğini sınırlandıran noktalara ise destek vektörleri denir. Verileri tam olarak ayırt eden hiper-düzlemin doğrusal olması durumuna doğrusal destek vektörleri denir. Eğitim için  $n$  sayıda örnekten oluşan eğitim kümesi aşağıdaki Denklem 3.1'de gösterildiği şekilde tanımlanmaktadır.

$$\{x^t, y^t\} \quad t = 1, 2, \dots, n, \quad y^t \in \{-1, +1\} \quad (3.1)$$

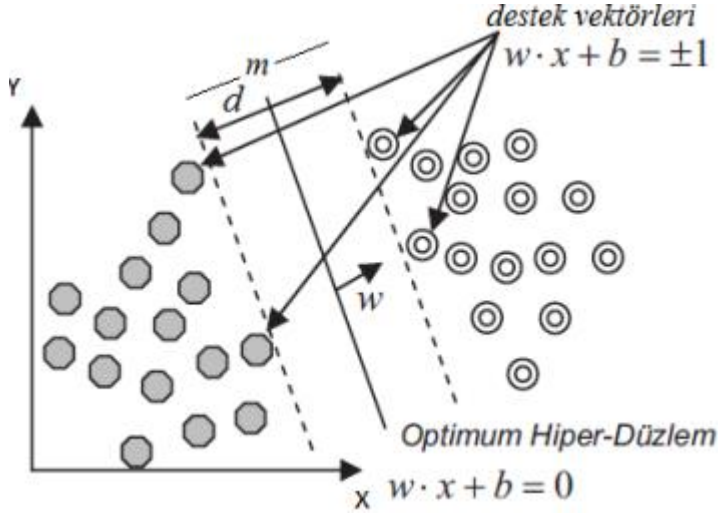
Optimum hiper düzleme ait eşitsizlikler Denklem 3.2 ve 3.3'de gösterildiği şekilde ifade edilmektedir.

$$wx^t + w_0 \geq +1 \quad , \quad y^t = +1 \quad (3.2)$$

$$wx^t + w_0 < -1 \quad , \quad y^t = -1 \quad (3.3)$$

Burada  $x \in R^N$  olup  $N$ -boyutlu bir uzayı,  $y \in \{-1, +1\}$  olup sınıf etiketlerini,  $w$  ağırlık vektörünü,  $w_0$  eğilim değerini ifade etmektedir.  $wx + w_0 = 0$  koşulunun sağlandığı durumdaki düzlem optimum hiperdüzlem olarak adlandırılmaktadır. Optimum hiper-düzlemin belirlenebilmesi için bu düzleme paralel ve sınırlarını oluşturacak iki hiper-düzlemin belirlenmesi gerekmektedir. Şekil 3.1'de gösterildiği üzere bu hiper-düzlemleri oluşturan noktalar destek vektörleri olarak adlandırılıp bu düzlemler  $wx^t + w_0 = \pm 1$  şeklinde ifade edilmektedirler. [45].





**Şekil 3.1:** Doğrusal olarak ayrılabilen veriler için hiper düzlem ve destek vektörleri [45].

Burada,  $w$  hiper-düzleme dik bir vektörü temsil etmektedir.  $m$  ise ayırım değerini ifade etmekte ve  $m = 2 \|w\|$  formülü ile hesaplanmaktadır. Amaç ayırım değerini maksimize etmek olduğundan  $\|w\|$  değerini minimize etmek gerekmektedir. En uygun hiper-düzlemi belirleyebilmek için aşağıdaki sınırlı optimizasyon probleminin çözümü gerekmektedir.

$r^t(w x^t + w_0) \geq +1$ , ve  $y^t \in \{-1, +1\}$  koşullarının sağlandığı durumda  $\min \left[ \frac{1}{2} \|w\|^2 \right]$ . Bu problem Lagrange denklemi ile çözülebilir. Bu işlem sonrasında Denklem 3.4 elde edilir.

$$L_P = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha^t [y^t (w x^t + w_0) + \sum_{t=1}^N \alpha^t] \quad (3.4)$$

Sonuçta doğrusal olarak ayrılabilen iki sınıflı bir örnek için karar fonksiyonu Denklem 3.5'deki gibi yazılmaktadır.

$$f(x) = \text{sign}(\sum_{t=1}^N \alpha^t \cdot y^t (x \cdot x^t) + w_0) \quad (3.5)$$

Bu denklem sıfırdan büyükse birinci sınıfa diğer hallerde ikinci sınıfa aittir şeklinde sınıflandırma yapılmaktadır.

### 3.1.1.2. Doğrusal Olmayan DVM

Giriş uzayında çekirdek fonksiyonları aracılığıyla doğrusal olmayan eğitim kümesinin yüksek boyutlu öznitelik uzayına dönüştürülerek, verilerin doğrusal olarak ayırt

edilebileceği düşünülür. Eğitim verilerinin bir kısmının optimum hiper-düzlemin diğer tarafında kalmasından kaynaklanan problem için bir  $\xi$  değişkeni tanımlanır. Sınırın maksimize edilmesi ve yanlış sınıflandırma hatalarının minimize edilmesi arasındaki denge,  $0-\infty$  arasında pozitif değerler alan ve  $C$  ile ifade edilen bir düzenleme parametresi tanımlanmasıyla kontrol edilebilmektedir [45]. Burada optimizasyon problemi Denklem 3.6'daki gibi gösterilmiştir.

$$f(x) = \min \left[ \frac{\|w\|^2}{2} + C \sum_{t=1}^r \xi_t \right] \quad (3.6)$$

Bu denklemin kısıtları ise Denklem 3.7'de ifade edilmiştir.

$$y^t((w \cdot x^t) + w_0) - 1 + \xi^t \geq 0 \text{ ve } \xi^t \geq 0 \text{ } t = 1, 2, \dots, n \quad (3.7)$$

### 3.1.1.3. Çok Sınıflı DVM

Bu kısma kadar olan işlemler iki sınıflı işlemler için uygundur. Çoklu sınıf için iki yaklaşım vardır. Bunlar Lagrange Fonksiyonu yaklaşımı ve bire karşı bir yaklaşımdır.

Lagrange Fonksiyonu Yaklaşımı; Lagrange fonksiyonları en küçükleme tipindeki problemi dual probleme çevirerek problemin çözülmesini sağlamaktadır.

Sınıflar arası Karşılaştırmalar ile Uygulanan Yaklaşımlar

Bire karşı bir yaklaşımı; sınıf sayısındaki her örnek kümesi diğer örnekler kümeleriyle ayrı ayrı eğitilmektedir.

Bire karşı hepsi yaklaşımı; her örnek kümesi, geri kalan tüm örnekler bir kümeye aitmiş gibi varsayılarak eğitilmektedir.

### 3.1.2. Doğrusal Diskriminant Analiz

Diskriminant analiz metodu sınıfları ayıran örüntü tanıma ve makine öğrenmesinde kullanılan istatistiksel bir metottur. Sınıf ayrımını kullanan yaklaşım Fisher tarafından bulunmuştur [46]. Sınıfların doğrusal olarak ayrılmasını amaçlamaktadır. Amacı aynı sınıfta olan verileri bir düzleme yansıtırken arasındaki mesafeyi minimum yapmaya çalışmaktır.

$$y = Ax \quad (3.8)$$

Denklem 3.8'de görüldüğü gibi  $A$  optimal yansıtma matrisi,  $x$  ise girdi değerini ifade etmektedir.  $A$  matrisi Denklem 3.9 ile hesaplanabilmektedir.

$$S_B W = \lambda S_W W \quad (3.9)$$

Denklem 3.9'da belirtilen  $S_W$ 'nin tersi alınabilmektedir. Eşitliğin iki tarafı da tersi ile çarpılarak Denklem 3.10 elde edilmektedir.

$$S_W^{-1} S_B W = \lambda W \quad (3.10)$$

$S_B$  değerinin alabileceği en yüksek değer sınıf sayısıdır. Bu durumda sınıf sayısı kadar 0 olmayan özdeğerler vardır denebilmektedir. Farklı veri setleri için her bir kümelenmenin ortalaması alınıp iki sınıf için kovaryans matrisi elde edilmektedir. Bütün kümelerin ortalaması alınıp kümedeki noktalar bu ortalamaдан çıkarılmakta ve bu değerlerin ortalaması alınarak sınıf içi dağılım kovaryans matrisi bulunmaktadır.

### 3.2. Diferansiyel Evrim Algoritması (DEA)

Optimizasyon bir sistemin istenen özelliklerini arttırırken istenmeyen özelliklerin azaltılması için yapılan işlemdir. Arama uzayında en iyi noktaya doğru olan her yaklaşım optimizasyonun bir parçasıdır.

Evrimsel algoritmalar optimizasyon türlerinden biridir. Evrimsel algoritmalar evrim ve doğal seçim kurallarına dayanarak doğadaki canlıların hayatta kalma mücadelelerini modelleyerek problem çözümlerinde kullanılır. Evrimsel algoritmalar türlerin doğal evriminden ilham alarak farklı alanlarda optimizasyon problemlerini çözmek için uygulanmıştır. Bu algoritmaların çoğunluğu en iyinin hayatta kalma prensibine göre çalışmaktadır. Her algoritmanın çözüme ulaştıran bir amaç fonksiyonu vardır. Bu amaca ulaşabilmek için doğadan ilham alınan mutasyon ve çaprazlama gibi fonksiyonları bulunmaktadır.

Evrimsel algoritmalar genelde rastgele bir çözüm ile işleme başlamaktadır. Bu sebeple bu algoritmaların genel özelliği hiçbir zaman çözüm garantisi sunmamasıdır. Bu algoritmalar için bulunan sonuç en iyi sonuçtur gibi kesin bir ifadeye yer verilmemektedir. Algoritmanın bulunduğu sonuç için en iyi çözüme yakın sonuç ifadesi kullanılabilir. Evrimsel algoritmalarından birisi de Diferansiyel Evrim Algoritmasıdır (DEA). DEA, Price ve Storn tarafından 1997 yılında sezgisel tabanlı bir optimizasyon algoritması olarak geliştirilmiştir [47]. Diferansiyel evrim algoritması diğer evrimsel algoritmalar gibi popülasyon üzerinden işlem yapmaktadır.

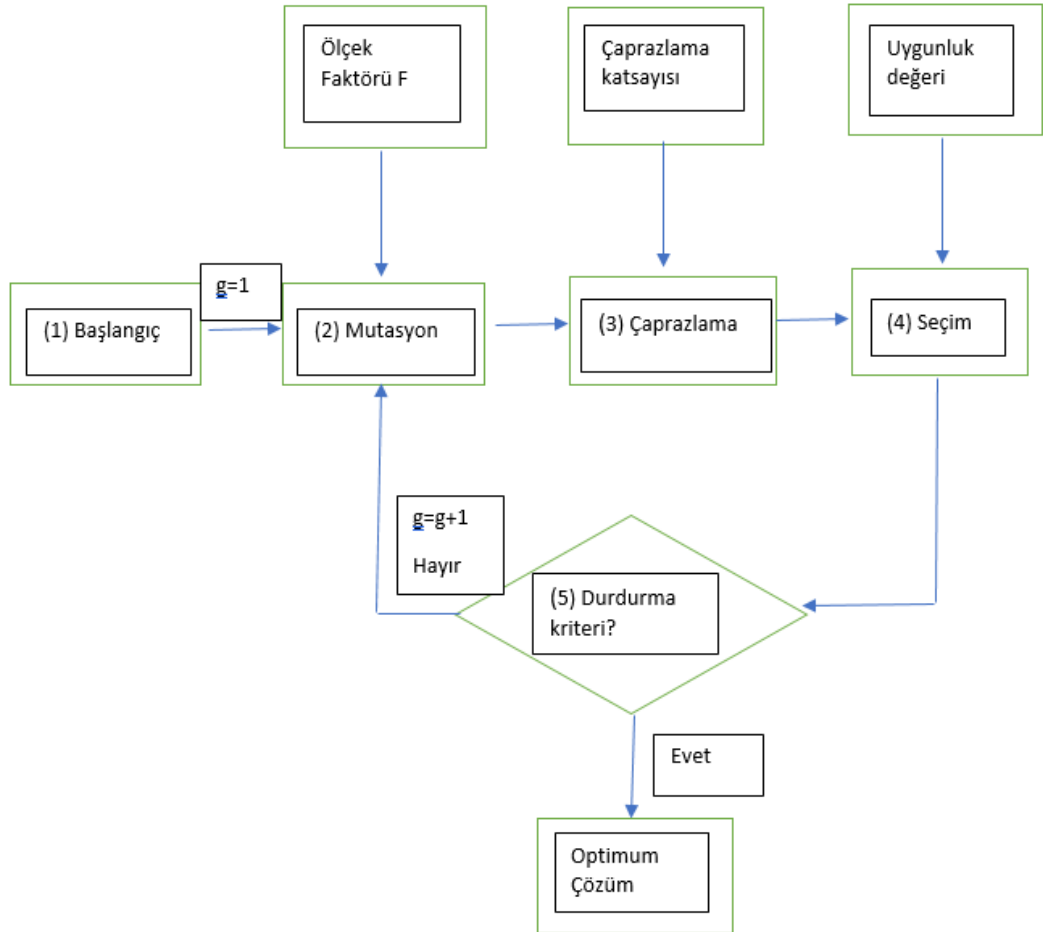
### 3.2.1. Popülasyon Yapısı

Popülasyon  $N_p$  adet bireyden oluşmaktadır. Popülasyon içerisinde ise her bir birey  $D$  gerçek değerli parametre içermektedir. Diferansiyel evrim algoritması  $N_p$  tane  $D$  boyutlu vektörden oluşmaktadır. Popülasyon Denklem 3.11'deki gibi tanımlanır.

$$P_{x,g} = (x_{i,g}), i = 0,1, \dots, N_{p-1} \quad g = 0,1, \dots, g_{max}$$

$$x_{i,g} = (x_{j,i,g}), j = 0,1, \dots, D - 1 \quad (3.11)$$

Denklem 3.11'de  $g = 0,1, \dots, g_{max}$  vektörün hangi nesle ait olduğunu göstermektedir.  $i = 0,1, \dots, N_{p-1}$  popülasyon içindeki bireyi,  $j = 0,1, \dots, D - 1$  birey içinde hangi parametrenin kullanıldığını belirtmektedir.



Şekil 3.2: Diferansiyel Evrim Algoritması Akış Şeması

Şekil 3.2'de gösterildiği üzere Diferansiyel Evrim Algoritması başlangıç parametrelerini belirleyerek başlamaktadır. Başlangıç popülasyonu için alt ve üst sınır

değerleri belirlenmektedir.  $D$  boyutlu  $b_j^L$  ve  $b_j^U$  olarak tanımlanan iki vektör için  $L$  alt sınırı  $U$  üst sınırı göstermektedir. Her vektörün parametresine Denklem 3.12'e göre alt ve üst sınırlara uygun parametreler atanır.

$$x_{j,i,g} = rand(0,1)(b_j^U - b_j^L) + b_j^L \quad (3.12)$$

Denklemdaki  $rand(0,1)$  ifadesi rastgele üretilen bir sayısı temsil etmektedir.

### 3.2.2 Mutasyon

Başlangıç parametreleri belirlenip popülasyon yapısı oluşturulduktan sonra rastgele seçilen vektörler için mutasyon işlemi uygulanmaktadır. Mutasyon işlemi Denklem 3.13'de ifade edildiği gibidir.

$$v_{i,g} = x_{r0,g} + F * (x_{r1,g} - x_{r2,g}) \quad (3.13)$$

Denklem 3.13'de ifade edilen  $r0$  baz olarak alınan vektör,  $r1$  ve  $r2$ 'de rastgele seçilmiş vektörleri ifade etmektedir. Baz vektörü ve rastgele seçilmiş vektörler birbirlerinden ve  $i$  'den farklı olmak zorundadırlar.  $F$  değeri ölçek faktörünü ifade etmekte ve  $[0,1)$  arasında bir değer almaktadır. Mutasyon işlemi ile ara popülasyon oluşturmaktadır. Bu yeni popülasyon Denklem 3.14'deki gibi ifade edilmektedir.

$$P_{v,g} = (v_{i,g}), i = 0,1, \dots, N_p - 1 \quad g = 0,1, \dots, g_{max}$$

$$v_{i,g} = (v_{j,i,g}), j = 0,1, \dots, D - 1 \quad (3.14)$$

Ara popülasyon  $P_{v,g}$ ,  $N_p$  adet mutant vektör  $v_{i,g}$  içermektedir.

### 3.2.3 Çaprazlama

Diferansiyel evrim operatörlerinden biri olan çaprazlama işlemi deneme vektörü oluşturmayı sağlamaktadır. Deneme vektörlerinin parametrelerinin mutant vektörlerden mi yoksa orijinal popülasyondan mı geleceği kararı verilmektedir.

$$u_{j,i,g} = \begin{cases} v_{j,i,g} , if (rand_j(0,1) \leq C_r \wedge j = j_{rand} \\ x_{j,i,g}, if (rand_j(0,1) > C_r \vee j \neq j_{rand} \end{cases} \quad (3.15)$$

Denklem 3.15'de gösterilen  $C_r$  çaprazlama olasılığını temsil etmektedir.  $C_r$ ,  $[0,1]$  arasında değer almaktadır. Denklemden ifade edildiği gibi eğer üretilen rastgele sayı çaprazlama faktöründen küçük ya da eşitse deneme faktörünün parametresi olarak

mutant vektörün parametresi alınmaktadır. Aksi durumda orijinal popülasyon parametresi alınmaktadır.

### 3.2.4. Seçim

Çaprazlama işlemi sonucu deneme vektörlerinin de oluşması ile birlikte diferansiyel evrim algoritması için seçim aşamasına geçilmektedir. Seçim yapılabilmesi için bir amaç fonksiyonu olmalıdır. Amaca en iyi hizmet eden nesil seçilip algoritma adımları devam etmektedir. Amaç fonksiyonu problemde farklılık göstermektedir. Örneğin; maliyet hesabı gibi işlemlerin optimizasyonunda en düşük maliyeti veren nesil tercih edilmektedir. Seçim işlemi sonrasında algoritmanın durdurma kriterinin sağlanıp sağlanmadığı kontrol edilmektedir. Eğer sağlanmışsa algoritma durdurulmakta değilse yeni nesile geçilmektedir. Amaç fonksiyonun belirli bir değere ulaşması ya da nesilin belirli bir sayıya ulaşması gibi kriterler kullanıcı tarafından algoritmanın durdurulması için problemin en başında belirlenmektedir.

### 3.2.5. Diferansiyel Evrim ile Öznitelik Seçimi

Bu tezde kullanılan DEA tabanlı öznitelik seçim metodu R.N. Khushaba ve arkadaşlarının yaklaşımıdır [48]. Bu yaklaşımın sözde kodu Şekil 3.3’de verilmiştir.

```
1 function [Err,Subset] = DEFS(data_tr,data_ts,DNF,PSIZE,Ld,classif,GEN)
2 degiskenlerin_tanımlanması()
3 H=öznitelik_sayisi;
4 NE=5;
5 NP=PSIZE=populasyon_sayısı=50(onerilen);
6 pop=populasyon_matrisi;
7 fit= hata_oranlarının_içinde_bulunduğu_matris
8 iBest=en_az_hata_oranlarını_içeren_öznitelik;
9 PD=error(x)<=error().mean();
10 ND=error(x)>error().mean();
11 populayon_oluşturma()
```

### Şekil 3.3: DEÖS’de değişkenlerin tanımlanması

Şekil 3.3’de gösterildiği üzere algoritma eğitim, test, seçilecek öznitelik sayısı, popülasyon büyüklüğü, sınıflandırma algoritması ve nesil sayısı parametrelerini kullanıcıdan alarak işleme başlamıştır. DE’nin üst ve alt sınır değişkenlerini tanımladığı gibi üst sınırı öznitelik sayısı olarak, alt sınırı 1 olarak kabul edilmiştir. Popülasyon büyüklüğü 50 olarak önerilmiştir. Uygulamada 50’den büyük rakam verildiğinde çalışma hızında yavaşlama gözlenmiştir. Pop, populasyon matrisini ifade

etmektedir. Amaç fonksiyonu pop matrisinde bulunan öznitelikler ile sınıflandırma sonucu en az hata oranının hesaplanması olarak belirlenmiştir. Sınıflandırma için LDA kullanılmıştır. Fit değeri 1-doğruluk oranı olarak hesaplanmıştır. iBest her nesilde en az hata oranını veren öznitelik alt kümelerini içeren matristir. PD değeri var olan hata ortalamasının bireysel hatadan büyük olması durumudur. ND ise var olan hata ortalamasının bireysel hatadan küçük olması durumudur.

```
1 degiskenlerin_tanimlanmasi()
2 x=rastgele_alt_kume()
3 hesapla=error(x)
4 fit(1)=hesapla
5 for g=2:GEN
6     v=mutant_vector(x) //mutant vektörleri
7     u=caprazlama(x,y) //deneme vektörleri
8     if(u<L && u>H) then
9         u()=x
10    else
11        u()=random()
12    endif
13    f=error(u)
14    if (error(u) <= error(current));
15        pop(:,j)=x
16        error(j)=f
17        if (error(u)<fit(iBest);
18            ibest=j
19        endif
20    endif
21 end
```

### Şekil 3.4: DEÖS'ün popülasyon ve hata oranının bulunması

DEÖS 'ün sonraki işlemleri Şekil 3.4'de gösterildiği adımlarda gerçekleşir;

- 1.satırda değişkenler tanımlanmıştır
- 2. satırlarda arasında rastgele değerler ile popülasyon oluşturulmuştur
- 3.satırda oluşturulan popülasyonun temsil ettiği özniteliklere göre sınıflandırma doğruluk oranı ile hata oranının hesaplanması yapılmıştır.
- 4.satırda fit değerine ilk alt kümenin sınıflandırma hata oranı eklenmiştir.
- 5. satırda iterasyon 1 artırılarak yeni jenerasyona geçilmiştir.
- 6.satırda mutant vektörler oluşturulmuştur.
- 7. satırda deneme vektörlerinin oluşumu gösterilmiştir. Mutant vektörlerden mi yoksa var olan popülasyon elemanlarından mı parametre seçileceği çaprazlama işlemi ile yapılmıştır. Çaprazlama işlemi sırasında DEA'sında sabit ölçek kullanılırken DEÖS'de değişen ölçek kullanılmıştır. Bunun etkisi, popülasyon

üyelerinin optimal çözümleri atlamadan sınırlar içinde hareket edebilmesine yardımcı olmaktadır [48]. Değişen ölçek faktörü Denklem 3.16'da gösterildiği gibi ifade edilmektedir.

$$F = \frac{c1 \cdot rand}{maksimum(x_{j,r1,g}, x_{j,r2,g})} \quad (3.16)$$

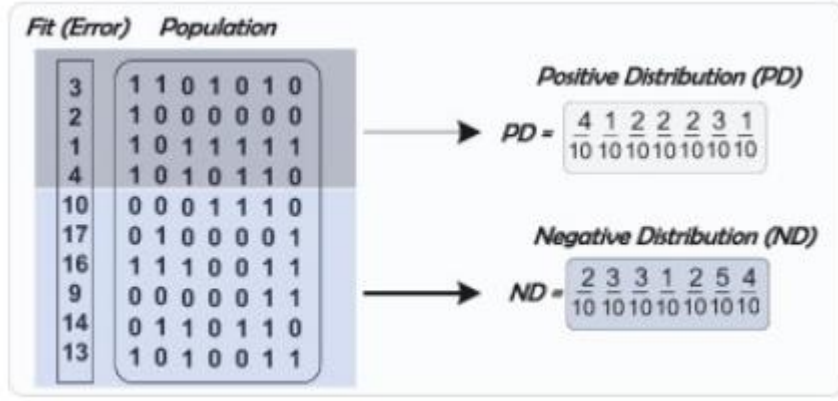
- 8-12 satırları arasında deneme matrisinin uygunluğu kontrol edilip, eğer H'den büyük ve L'den küçük ya da daha önce oluşturulan vektörlerle aynı ise değiştirilmiştir. Rastgele yeni deneme matrisi oluşturulmuştur.
- 13'de deneme matrisinin hata oranı hesaplanmıştır.
- 14-20 satırları arasında deneme matrisinin hata oranı düşükse popülasyon matrisi yerine deneme matrisi eklenmiştir. Değilse var olan matrisle işleme devam edilmiştir.
- 21'de durdurma kriteri sağlanmadıysa sonraki jenerasyona geçilmiştir.

Bu algoritmanın avantajı gerçek değerlerin kullanılabilir oluşudur. Fakat bu avantaj aynı zamanda bir zorluğu da beraberinde getirmiştir. Bu zorluk da şudur; en iyi öznitelik alt kümesinin seçimi sırasında sayılar küsuratlı gelebilmekte ve bu sayıların yuvarlanması gerekmektedir. Küsuratlı elde edilen sayıların farklı olacağı kesindir ama yuvarlanınca oluşan yeni sayılar eşsiz olacağının garantisi yoktur. Bu probleme çözüm olarak rulet tekerleği ağırlık şeması kullanılmıştır [48]. Bu şemada her bir özelliğin olasılığının kendisi ile ilgili bir maliyet hesabı uygulanmıştır. Dağıtım faktörü  $FD$  'nin hesaplanması  $g$  iterasyonu,  $j$  özneliği belirtmek üzere Denklem 3.17'deki gibi hesaplanmıştır.

$$FD_{j,g} = a_1 \cdot \left( \frac{PD_j}{PD_j + ND_j} \right) + \frac{NF - DNF}{NF} \cdot \left( 1 - \frac{PD_j + ND_j}{mak(PD_j + ND_j)} \right) \quad (3.17)$$

$PD$  pozitif dağılım,  $ND$  negatif dağılımı ifade etmektedir. Değerleri ise Şekil 3.5 [48]'deki gibi bulunmaktadır. Popülasyonların hata oranları belirlenmekte, belirlenen değerden küçük olanlar pozitif, büyükler negatif dağılımı temsil etmektedir. Popülasyondaki her bir vektörün hata değerinin popülasyon sayısına bölümü dağılımı vermektedir.





**Şekil 3.5:** Dağılım faktörünün hesaplanması [48].

Bu tezde iki farklı veri seti için değişen parametreler aşağıdaki gibi belirlenmiştir;

- Bonn Üniversitesi için;
  - Veri boyutu 100 x 24 (örnek x öznelik) olmaktadır. En az 1 özellik en çok bütün özellikler kullanılabilir şekilde olduğundan, L=1, H=24 olarak belirlenmiştir.
- BCI verisi için;
  - Veri boyutu 180 x 432(örnek x özellik) olmaktadır. L=1, H =432 olarak belirlenmiştir.

## 4. BULGULAR

Bu tezde EEG veri kümeleri üzerinde sınıflandırma işlemi ile en yüksek doğruluk oranı elde edilmeye çalışılmıştır. Bu sebeple farklı örüntü tanıma yöntemleri ve öznitelik seçim metotları uygulanmıştır. Kullanılan iki farklı veri setinden farklı sonuçlar elde edilmiştir. Bu sebeple bulgular iki başlık altında toplanmıştır.

### 4.1. Bonn Verisi İçin Bulgular

İlk veri seti Bonn Üniversitesi'nin epilepsi hastalığı için toplanmış verisidir. Bu sette 5 veri kümesi bulunmaktadır. Kümeler A, B, C, D ve E olmak üzere 5 adettir. Verinin genel özeti; A ve B kümeleri sağlıklı ve A'nın gözü açık B'nin gözü kapalı C, D ve E kümeleri ise hastadır. E kümesi hastanın nöbet anında alınmış verileridir. C, D ise nöbet halinde değil ama hasta olan deneklerden alınmış verilerdir. C ve D'nin birbirinden farkı ise elektrotların bağlandığı kısımların farklı olmasıdır.

İlk olarak veri kümesine bant geçiren filtresi uygulanmıştır. Bant geçiren filtresi yüksek ve alçak bant frekansları dışında kalan alanları bastıran ve sadece o arada kalan verileri ortaya çıkaran bir yöntemdir. 0,53 Hz'den küçük frekanslı sinyalleri ve 40 Hz'den büyük frekanstaki sinyalleri bastırmıştır.

Filtrelenmiş veriye ADD ve 1D-YİÖ uygulanarak öznitelik çıkarımı yapılmıştır. Elde edilen öznitelik matrisine TBA uygulanarak boyut küçültmeye gidilmiş ve farklı sayıda TBA parçası seçilerek sınıflandırılmıştır. Öznitelik matrisine DEÖS ile öznitelik seçimi uygulanmış ve yüksek sınıflandırma doğruluk oranı veren öznitelik sütunlarıyla yeni alt kümeler oluşturulmuştur. Elde edilen alt kümeler sınıflandırıcılar ile test edilmiştir. Sınıflandırma doğruluk oranları karşılaştırılmıştır.

Kısaca özetlemek gerekirse Bonn Üniversitesi veri seti için sırasıyla denenen işlemlerin özeti şöyledir;

- Filtreleme -> ADD -> LDA\DVM
- Filtreleme -> ADD ->DEÖS -> LDA\DVM
- Filtreleme -> 1D-YİÖ ->LDA\DVM

- Filtreleme -> ADD -> TBA -> LDA

Literatürde yapılan çalışmalar ve Bonn verisinin farklı kümeleri varlığında elde edilen sınıflandırma doğruluk oranları özet tablosu Tablo 4.1’de gösterilmektedir. Bu tabloda da gösterildiği üzere veri setindeki farklı sınıflar (A, B, C, D, E) ikili (AxB, AxC, AxD, AxE) ve üçlü (AxBxC, AxBxD, AxBxE, BxCxD, BxCxE, CxDxE, AxCxD, AxCxE, BxDxE) ve farklı ikili ve üçlü gruplar (ABxCDxE, ABCDxE) olmak üzere sınıflandırılmaya çalışılmış ve doğruluk oranları kıyaslanmıştır. Yapılan işlemler için iki sınıflandırıcı kullanılmıştır. İlk olarak LDA ile elde edilen sonuçları incelenmiştir.

İkili gruplar halinde sınıflandırma işleminde A ve B sınıflarının her biri 100x24 boyutunda kümeleri temsil etmektedir. Sınıf etiketleri eklenerek 100x25 boyutunda matrisler oluşmuştur. AxB ile 200x25 boyutunda 2 sınıflı bir matris elde edilmiştir.

Üçlü gruplar halinde sınıflandırma işleminde A, B ve C sınıflarının her biri 100x24 boyutunda kümeleri temsil etmektedir. Sınıf etiketleri eklenerek 100x25 boyutunda matrisler oluşmuştur. AxBxC ile 300x25 boyutunda 3 sınıflı bir matris elde edilmiştir.

Farklı üçlü gruplar halinde sınıflandırma işleminde A, B, C, D ve E sınıflarının her biri 100x24 boyutunda kümeleri temsil etmektedir. Sınıf etiketleri eklenerek 100x25 boyutunda matrisler oluşmuştur. ABxCDxE ile 500x25 boyutunda 3 sınıflı bir matris elde edilmiştir. AB 200x25 boyutunda 1 sınıflı bir kümeyi, CD 200x25 boyutunda 1 sınıflı bir kümeyi ve E 100x25 boyutunda 1 sınıflı bir kümeyi temsil etmiştir.

**Tablo 4.1:** Literatürde Bonn verisinin farklı kümeleri ile yapılan çalışmalar ve sınıflandırma doğruluk oranları

Çalışma Sahibi	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk oranı (%)
Z. Lasefr v.d	ADD	AxE	YSA	98,1
		AxE	DVM	96,2
Md. Rasih v.d.	ADD (db4)	AxE	Sinir ağları	100
		BxE		99,5
		CxE		99,5
		DxE		99,0
		AxCxE		80,0
		BxCxE		78,7
		AxDxE		80,0
		BxDxE		79,3
A. Ahmadi v.d.	ADD (db4, db2, bior)	AxE	DVM	99,67
		BxE		98,44
		CxE		98,25
		DxE		95,15
		ABCDxE		97,85
W. Ren v.d	ADD/ DPD/ SampEn +TBA/Fisher skor/mRMR	AxDxE	Karar ağacı	98,95
S. Ramakrishnan v.d	ADD (4. Seviye)	AxE	DVM	100
		ABCDx E		99
		AxDxE		98
		ABxCDxE		96
		AxBxCxDxE		95
O. Salem v.d.	ADD (4. seviye)	ABCDxE	Karıncı kolonisi sınıflandırması	ROC eğrisi Hata oranı= 9
S. Bose v.d.	ADD (5. seviye)	AxDxE	Karar Ağaçları	98
D. Chen v.d.	ADD	ABCDxE	DVM	99,33
U.R. Archarya v.d	DPD (rbio6.8)	AxE	Karar Ağacı	91,3
			Bulanık	96,7
			k-en yakın komşu	93,3
			Naif Bayes	82
Y.Kaya v.d	1D-YiÖ	AxE	DVM	98,5
		AxE	YSA	98
		AxD	DVM	99
		AxD	YSA	98,5
		DxE	DVM	93
		DxE	YSA	93,5
		ABxCDE	DVM	92,8
		ABxCDE	YSA	95,40
		AxDxE	DVM	94
		AxDxE	YSA	94,33

#### 4.1.1.LDA ile Sınıflandırma Doğruluk Oranı

Sınıflandırıcı olarak LDA kullanılan deneyler bu başlık altında incelenmiştir. LDA ile sınıflandırmada doğruluk oranı 10'lu çapraz doğrulama ile bulunmuştur. Deney sonucunda ortalama doğruluk oranı ve standart sapma değerleri ilgili tablolara eklenmiştir.

##### 4.1.1.1.ADD Kullanılarak Öznitelik Matrisi Oluşturma

Filtrelenen veriye ADD uygulanmıştır. EEG verisine ADD uygulanması yüksek geçiren ve alçak geçiren filtre ile alt bant katsayıları elde edilmesini sağlar. ADD uygulanırken veri tek kanallı olduğu için tek bir kanaldan alt kanallar elde edilmiştir. Dönüştürülen veri 100x4096 boyutundadır. 4096 bir sinyalin uzunluğunu ifade ederken 100 örnek sayısını belirtmektedir. 4. dereceden ADD uygulanması sonucu bir kanallı bir sinyalden A4, D4, D3, D2 ve D1 alt bantları elde edilir. D3, D4 ve A4 alt bantları seçilmiştir, çünkü bu alt bant frekansları veri için ilgili aralığı temsil etmektedir. Böylece bu aralıktan D3, D4 ve A4 için alt bant katsayıları cD3, cD4 ve cA4 elde edilmiştir. Bu durumda bir kanallı veri 3 alt banda bölünmüştür. Her bir alt bandın katsayıları yardımıyla 8 adet istatistiksel özellik çıkarılmıştır. Böylece veri setinin her alt bandı için çıkarılan özellikler maksimum, minimum, standart sapma, ortalama, mutlak toplam, basıklık (kurtosis), çarpıklık (skewness), ve enerjidir. Her bir sınıf için (A, B, C, D ve E) 100x24 boyutlu bir matris oluşturmuştur. Etiket verisi de eklenince bütün sınıflar ile birlikte (100x5=500) verinin boyutu 500x25 olmuştur.

Bu tezde Tablo 4.1'de görüldüğü gibi ikili ve üçlü olarak sınıfların sınıflandırma doğruluk oranları kıyaslanmıştır. Kıyaslanan tablolar aşağıdaki gibidir. A, B, C, D ve E sınıflarının veri setleri kıyaslanması Tablo 4.2'de verilmiştir. Bu ikili kümelerin sınıflandırmasında 2 sınıf için 2 farklı etiket kullanılmış ve sonuçlar Tablo 4.2, 4.3, 4.4'e eklenmiştir.

Tablo 4.2'de literatürde de oldukça yaygın görülen etiketli A sınıfının (100x25) varlığında diğer sınıflar ile yaptığı ikili kümeler sonucu oluşan kümelerin (200x25) LDA ile sınıflandırma doğruluk oranının sonuçları gösterilmiştir.

**Tablo 4.2:** İkili olarak A sınıfı ile diğer sınıfların varlığında LDA ile doğruluk oranı kıyaslanması.

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	ADD (4. Seviye, db2)	AxB	LDA	$0,915 \pm 0,0634$
		AxC		$0,97 \pm 0,0331$
		AxD		$0,9399 \pm 0,0538$
		AxE		$0,9799 \pm 0,0244$

Tablo 4.2’deki doğruluk oranlarına göre en farklı sınıfların oluşturduğu kümelerin (sağlıklı ve sağlıksız) doğruluk oranları benzer kümelere göre (sağlıklı x sağlıklı) daha yüksek çıkmıştır. En yüksek doğruluk oranı AxE sınıfları ile oluşturulan kümedir. Çünkü A sağlıklı E ise hasta ve nöbet halinde olan kişilerden alınmış verilerdir.

Tablo 4.3 ‘de B sınıfının diğer sınıflarla ikili oluşturduğu veri kümeleri ile sınıflandırma doğruluk oranları kıyaslanmıştır. Sınıflandırıcı olarak LDA kullanılmıştır.

**Tablo 4.3:** İkili olarak B sınıfı ile diğer sınıfların varlığında LDA ile doğruluk oranı kıyaslanması.

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	ADD (4. Seviye, db2)	BxC	LDA	$0,975 \pm 0,0335$
		BxD		$0,98 \pm 0,0399$
		BxE		$0,98 \pm 0,0244$

Tablo 4.3’ de gösterildiği üzere tıpkı A kümesi varlığında kıyaslamalarda olduğu gibi oldukça yüksek ve benzer sonuçlar elde edilmiştir. Bu durumun sebebi de A ve B sınıflarının birbirinden tek farkının göz açıklığı olmasıdır. Tablo 4.4’de C ve D sınıflarının ikili küme oluşturması sonucu sınıflandırma doğruluk oranlarının LDA ile bulunması gösterilmiştir.

**Tablo 4.4:** İkili olarak C ve D sınıfları ile diğer sınıfların varlığında LDA ile doğruluk oranı kıyaslanması.

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kullanılan Küme	Sınıflandırma	Doğruluk Oranı
Bonn	ADD (4. Seviye, db2)	CxD	LDA	$0,735 \pm 0,0743$
		CxE		$0,975 \pm 0,0335$
		DxE		$0,940 \pm 0,0489$

C ve D kümelerinin hem hasta olmaları hem de nöbet halinde bulunmaması sebebiyle ortak özellikleri vardır. Bu sebeple Tablo 4.4’de en düşük doğruluk oranını CxD kümesiyle elde edildiği gösterilmiştir. C ve D kümelerinin tek farkının elektrotların bağlandığı yer olması sebebiyle literatürde de çok fazla üzerinde deney yapılan kümeler değillerdir.

Üçlü kombinler halindeki sınıfların varlığında sınıflandırma doğruluk oranı kıyaslanması her bir sınıfa ayrı ayrı etiket verilerek yapılmıştır. A, B, C gibi her bir sınıfın etiketi ayrı ayrı verilir ve sınıflandırmaya tabi tutulur. A sınıfı 100x25 iken AxBxC kümesi 300x25’dir. Elde edilen doğruluk oranları Tablo 4.5’de gösterildiği gibidir.

**Tablo 4.5:** Üçlü kombinler halinde sınıfların varlığında LDA ile doğruluk oranı kıyaslanması

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümelere	Sınıflandırıcı	Doğruluk Oranı
Bonn	ADD (4. Seviye, db2)	AxBxC	LDA	0,9133 ± 0,0520
		AxBxD		0,90 ± 0,0516
		AxBxE		0,8766 ± 0,0422
		AxDxE		0,8766 ± 0,063
		AxCxE		0,88 ± 0,059
		BxDxE		0,94 ± 0,032
		BxCxE		0,963 ± 0,0406

Tablo 4.5’de gösterildiği gibi içerisinde benzer sınıfların bulunduğu kümelerin sınıflandırma doğruluk oranları diğer kümelere göre düşük çıkmaktadır. En farklı sınıfların birleşmesi ile oluşmuş olan BxDxE ve BxCxE kümelerine göre AxBxC gibi yakın sınıfların bulunduğu kümeler daha düşük değerler vermiştir.

Bu kümelerin haricinde literatürde de özellikle tercih edilen [13] farklı kümelemelerle de denenmiştir. AxBxCxDxE, ABCDxE, ABxCDxE gibi kombinlerde test edilmiştir. AxBxCxDxE bu veri kümesinde her bir örnek ayrı etiketlenmiştir. ABCDxE ise veri kümesinde ABCD bir bütün olarak tek bir sınıfı, E sınıfı başka bir sınıfı temsil edecek şekilde etiketlenmiştir. ABxCDxE verisinde AB bir sınıfı, CD bir sınıfı, E ise öteki sınıfı temsil etmiştir. Farklı kombinlerle sınıf oluşturma ve doğruluk oranlarının kıyaslanması Tablo 4.6’daki gibidir.

**Tablo 4.6:** Farklı sınıflar arasında LDA ile doğruluk oranı

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	ADD (4. Seviye, db2)	AxBxCxDxE	LDA	0,722 ± 0,0509
		ABCDxE		0,968 ± 0,027
		ABxCDxE		0,932 ± 0,0411

Tablo 4.6’da gösterildiği şekilde sınıfların birleştirilmesi sonucu iki sınıf elde edilmiş olan ABCDxE kümesinin doğruluk oranı daha yüksek çıkmıştır. Sebebi bütün sınıflara karşı nöbet hali gibi fiziksel olarak bile ayırt edilebilen bir sınıf ile kümelenmesidir. En düşük sonuç AxBxCxDxE kümesinin olmuştur. Sebebi beş farklı sınıfın ayrı ayrı kümelenmeye çalışmasıdır.

Bu sonuçlara göre hem ikili hem üçlü kombinlerle yapılan kümelerin bulunduğu hem de literatürde farklı kombinlerle denenen sınıfların birleşimiyle oluşmuş veri kümelerinin sınıflandırma doğruluk oranları oldukça yüksek çıkmıştır. Bu sebeple diferansiyel evrim algoritması ile öznitelik alt kümesi belirleyebilmek için daha düşük doğruluk oranı veren kümelere ihtiyaç vardır. Bunun için AxBxCxDxE kümesi tercih edilmiştir.

#### 4.1.1.2. DEÖS Kullanılarak Öznitelik Oluşturma

Diferansiyel Evrim ile Öznitelik Seçimi, ADD ile oluşturulmuş öznitelik kümesine uygulanmıştır. Etiketli verinin boyutu 500x25’dir. Doğruluk oranı düşük olduğu için seçilen 5 sınıflı veri kümesine diferansiyel evrim algoritmasının farklı parametreleri ile denemeler yapılmıştır. Çaprazlama oranı (CR) parametresi için 0.3 - 0.7 değerleri, iterasyon sayısı (g) değeri için ise 10000, 1000 ve 100 değerini alarak hesaplamalar yapılmıştır. Bu değerler farklı öznitelik alt küme sayılarıyla hesaplanmış ve sonuçlar LDA ile test edilmiştir. Tablo 4.7’de 24 tane sütunu bulunan öznitelik kümesinden DEÖS ile 10 adet öznitelik seçilmesi sonucu oluşan alt kümenin doğruluk oranı gösterilmiştir. DEÖS algoritmasının doğruluk oranını hesaplarken aldığı parametreler Tablo 4.7’ye eklenmiştir.



**Tablo 4.7:** 10 özniteliğe sahip alt kümenin LDA ile sınıflandırma doğruluk oranı gösterimi

CR	İterasyon sayısı	Sınıflandırıcı	Doğruluk Oranı
0,3	10000	LDA	$0,6819 \pm 0,0568$
0,5	10000		$0,7024 \pm 0,064$
0,7	10000		$0,7020 \pm 0,0641$
0,3	1000		$0,716 \pm 0,0691$
0,5	1000		$0,6819 \pm 0,0568$
0,7	1000		$0,6559 \pm 0,0731$
0,3	100		$0,716 \pm 0,0691$
0,5	100		$0,722 \pm 0,0689$
0,7	100		$0,6900 \pm 0,0531$

Tablo 4.7’de gösterildiği üzere iterasyon sayısının 100 ve 10000 olduğu durumlarda en yüksek doğruluk oranı çaprazlama oranının 0,5 olduğu durumda bulunmuştur. Tablo 4.8’de 24 öznitelikten DEÖS algoritmasıyla 15 elemanlı öznitelik kümesi elde edilip doğruluk oranı kıyaslaması yapılmıştır.

**Tablo 4.8:** 15 özniteliğe sahip alt kümenin LDA ile sınıflandırma doğruluk oranı gösterimi

CR	İterasyon sayısı	Sınıflandırıcı	Doğruluk Oranı
0,3	10000	LDA	$0,73 \pm 0,0627$
0,5	10000		$0,73 \pm 0,0538$
0,7	10000		$0,74 \pm 0,0551$
0,3	1000		$0,722 \pm 0,0524$
0,5	1000		$0,73 \pm 0,0507$
0,7	1000		$0,73 \pm 0,0627$
0,3	100		$0,722 \pm 0,0524$
0,5	100		$0,722 \pm 0,0524$
0,7	100		$0,6519 \pm 0,0581$

Tablo 4.8’de gösterildiği üzere en yüksek doğruluk oranı 0,7 çaprazlama değeri ve 10000 iterasyon sayısı ile %74’dür. Tabloya bakıldığında 1000 iterasyonda da en yüksek değer 0,7 çaprazlama oranı ile elde edildiği gösterilmiştir. Buna rağmen 100 iterasyonda en yüksek değere 0,7 çaprazlama oranı değeri ile ulaşamamıştır. Bunun sebebi de 100 iterasyonun en iyi sınıflandırma doğruluk oranına sahip alt kümeyi seçemeyecek kadar az olmasıdır.

24 öznitelikten DEÖS algoritmasına 20 elemanlı öznitelik kümesi elde edilip doğruluk oranı kıyaslaması yapılmıştır. DEÖS ile doğruluk oranı hesaplanırken aldığı parametreler Tablo 4.9'a eklenmiştir.

**Tablo 4.9:** 20 özniteliğe sahip alt kümenin LDA ile sınıflandırma doğruluk oranı gösterimi

CR	İterasyon sayısı	Sınıflandırıcı	Doğruluk Oranı
0,3	10000	LDA	0,734 ± 0,0506
0,5	10000		0,734 ± 0,0506
0,7	10000		0,734 ± 0,0506
0,3	1000		0,734 ± 0,0506
0,5	1000		0,734 ± 0,0506
0,7	1000		0,734 ± 0,0506
0,3	100		0,728 ± 0,0506
0,5	100		0,734 ± 0,0506
0,7	100		0,728 ± 0,0506

Tablo 4.9'da hem çaprazlama oranı hem de iterasyon sayısının değişmesine rağmen doğruluk oranının pek değişmemesinin sebebi 24 öznitelikli kümeden 20 öznitelikli alt küme oluşturmaya çalışmasıdır. Alt küme oluşturabileceği eleman sayısının sınırlı olması bu sonucu oluşturmuştur.

#### 4.1.1.3. TBA Kullanılarak Öznitelik Çıkarımı

ADD uygulanarak oluşturulmuş veri kümesine TBA yöntemi uygulanmıştır. Sınıflandırma doğruluk oranını arttırmak için denenen boyut azaltma yöntemlerinden TBA ile farklı parça sayıları ile alınan sonuçlar Tablo 4.10'daki gibidir. TBA sonucu elde edilen parçalarla sınıflandırma işlemi LDA ile 10 kat çaprazlama yöntemi ile hesaplanmıştır.

**Tablo 4.10:** TBA ile sınıflandırma doğruluk oranlarının kıyaslanması

TBA bileşen sayısı	Doğruluk Oranı
10	0,7669 ± 0,1525
15	0,722 ± 0,1852
20	0,7319 ± 0,1675

Tablo 4.10’da gösterildiği üzere en yüksek doğruluk oranı 24 öznitelikli kümeden 10 bileşen oluşturarak %76 olarak bulunmuştur.

#### 4.1.1.4. 1D-YİÖ Kullanılarak Öznitelik Oluşturma

Bonn verisinde 1D-YİÖ yöntemi kullanıldığında farklı doğruluk oranları elde edilmiştir. Bu doğruluk oranlarının kıyaslaması ilk olarak A sınıfı ile ikili olarak diğer sınıfların kombinlenmesi ile elde edilen kümelerde denenmiştir. İkili olarak A sınıfı ile diğer sınıfların varlığında doğruluk oranının kıyaslanması Tablo 4.11’de gösterilmiştir.

**Tablo 4.11:** A sınıfı ile diğer sınıfların ikili oluşturdukları kümenin LDA ile doğruluk oranı

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	1D-YİÖ	AxB	LDA	0,89 ± 0,0634
		AxC		0,880 ± 0,0509
		AxD		0,9349 ± 0,0449
		AxE		0,885 ± 0,0867

Tablo 4.11’de gösterildiği üzere en düşük doğruluk oranı AxC kümesinde oluşmuştur. En yüksek doğruluk oranı ise AxD kümesi ile elde edilmiştir.

**Tablo 4.12:** B sınıfı ile diğer sınıfların ikili oluşturdukları kümenin LDA ile doğruluk oranı

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	1D-YİÖ	BxC	LDA	0,9499 ± 0,0634
		BxD		0,9799 ± 0,0509
		BxE		0,795 ± 0,0449

Tablo 4.12’de en yüksek sınıflandırma doğruluk oranı BxD ile bulunmuştur. Tablo 4.11’deki sonuçlarda da A kümesinin D ile yaptığı birleşim sonucu en yüksek doğruluk oranı bulunmuştur. Bu benzerliğin sebebi A ve B kümelerinin oldukça benzer deneklerin verileriyle oluşturulmasıdır.

C ve D sınıflarının diğer sınıflarlar varlığında ayrı ayrı ikili kombin kümelerinin LDA ile doğruluk oranlarının kıyaslanması Tablo 4.13’de gösterilmiştir

**Tablo 4.13:** C ve D sınıflarının diğer sınıflar varlığında ayrı ayrı ikili kombinler ile oluşturdukları kümenin LDA ile doğruluk oranı

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	1D-YİÖ	CxD	LDA	0,61 ± 0,0634
		CxE		0,7649 ± 0,0509
		DxE		0,675 ± 0,0449

Tablo 4.13’de gösterildiği üzere en yüksek doğruluk oranı E sınıfı ile yapılan kümelerle oluşmuştur. Sebebi ise E kümesinin epilepsi nöbeti halinde oluşturulmasıdır. A, B, C, D ve E sınıfları ile oluşturulan ikili doğruluk oranları kıyaslamadan sonra üçlü kombinler oluşturulup doğruluk oranları DVM ve LDA ile kıyaslanmıştır. Tablo 4.14’de üçlü kombinlerin 10’lu çaprazlama ile doğruluk oranı hesaplanmıştır.

**Tablo 4.14:** Üçlü kombinler halinde kümelerin sınıflandırma doğruluk oranlarının LDA ile ölçülmesi.

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	1D-YİÖ	AxBxC	LDA	0,8466 ± 0,0618
		AxBxD		0,86 ± 0,0879
		AxBxE		0,736 ± 0,0887
		AxCxD		0,676 ± 0,065
		AxCxE		0,7366 ± 0,0737
		AxDxE		0,666 ± 0,0577
		BxCxD		0,75 ± 0,05
		BxCxE		0,7266 ± 0,0573
		BxDxE		0,63 ± 0,0959
		CxDxE		0,526 ± 0,064

Tablo 4.14’de gösterildiği üzere en düşük doğruluk oranına sahip üçlü kümelene hepsi hasta deneklerden toplanan veri kümeleri olan CxDxE kümesidir.

İkili ve üçlü kombinlerin ardından dördü kombinler oluşturulup sınıflandırma doğruluk oranları kıyaslanmıştır. Literatürde Bonn verisinin sınıflarını dördü olarak

kombinleyen başka bir çalışma bulunmamaktadır. Dörtlü kombinlerin LDA ile doğruluk oranlarının ölçümü Tablo 4.15’de gösterilmiştir.

**Tablo 4.15:** Dörtlü kombinlerin LDA ile doğruluk oranlarının kıyaslanması

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı
Bonn	1D-YiÖ	AxBxCxD	LDA	0,71 ± 0,0618
		AxBxCxE		0,6775 ± 00575
		AxBxCxE		0,6799 ± 0,0857

#### 4.1.2. DVM ile Sınıflandırma Doğruluk Oranı

Sınıflandırıcı olarak DVM kullanılan deneyler bu başlık altında incelenmiştir. DVM ile sınıflandırmada doğruluk oranı 10’lu çapraz doğrulama ile bulunmuştur.

##### 4.1.2.1. ADD Kullanılarak Öznitelik Oluşturma

Bu kısımda ADD sonucu oluşturulan öznitelikler ile işlemler yapılmıştır. Oluşan öznitelik kümeleri ikili, üçlü, dörtlü ve beşli olarak kombinlenerek sınıflandırma doğruluk oranları hesaplanmıştır. Tablo 4.16’daki gibi A sınıfının varlığında diğer sınıflar ile ikili kümeler oluşturması sonucu gösterilmiştir. Bu sonuçları veren DVM parametreleri tabloya eklenmiştir.

**Tablo 4.16:** İkili olarak A sınıfı ile diğer sınıfların varlığında DVM ile doğruluk oranları kıyaslanması.

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kullanılan Küme	Sınıflandırıcı	Doğruluk Oranı	Parametreler
					c / gamma /kernel
Bonn	ADD (4. Seviye, db2)	AxB	DVM	0,9125	1 /0,15/ rbf
		AxC		0,95	1 /0,1/ rbf
		AxD		0,975	10 /0,1/ rbf
		AxE		0,9875	1/0,1/ rbf

Tablo 4.16’da beklenildiği üzere en yüksek sonucu veren kümelenme sağlıklı ve epilepsi nöbeti halindeki bireylerin verilerini içeren AxE sınıflarından oluşmaktadır. En düşük doğruluk oranı ise birbirine oldukça benzeyen iki sağlıklı bireyin verilerini içeren AxB sınıflarıdır. Tablo 4.17’de B sınıfı varlığında diğer sınıflar ile doğruluk oranı kıyaslaması yapılmıştır.

**Tablo 4.17:** İkili olarak B sınıfı ile diğer sınıfların varlığında DVM ile doğruluk oranı kıyaslanması.

Kullanılan Veri	Öznelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı	Parametreler
					c/ gamma/kernel
Bonn	ADD (4. Seviye, db2)	BxC	DVM	0,9812	1/ 0,1/ rbf
		BxD		0,9687	1/ 0,1/ rbf
		BxE		0,9812	10/ 0,1/ rbf

Tablo 4.17’de gösterildiği üzere kendisine çok benzeyen A ile bir küme oluşturulmadığı için doğruluk oranları oldukça yüksek çıkmıştır. En yüksek doğruluk oranlarından birinin E sınıfı ile olması E’nin nöbet halindeki kişilerin verilerini içeren küme olmasındandır. Tablo 4.18’de C ve D sınıfları ile diğer sınıfların ikili oluşturdukları kümelerin sınıflandırılma sonucu doğruluk oranları gösterilmiştir.

**Tablo 4.18:** İkili olarak C ve D sınıfları ile diğer sınıfların varlığında DVM ile doğruluk oranı kıyaslanması.

Kullanılan Veri	Öznelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı	Parametreler
					c/ gamma/ kernel
Bonn	ADD (4. Seviye, db2)	CxD	DVM	0,7625	10/ 0,125/ rbf
		CxE		0,9625	1/ 0,1 /rbf
		DxE		0,9625	10/ 0,1/ rbf

Tablo 4.18’de gösterilen sonuçlara göre en düşük doğruluk oranı, sınıfları arasında benzer özelliklerin olması sebebiyle CxD kümesidir. Diğer sınıfların kıyaslanması nöbet halindeki hastadan alınmış olması sebebiyle sınıflar arasında belirgin fark oluşturarak doğruluk oranını arttırmıştır. Tablo 4.19’da üçlü kümelerin sınıflandırması sonucu en iyi sonuçlar ve en iyi sonuçları veren parametreler gösterilmiştir.

**Tablo 4.19:** Üçlü kombinler halindeki sınıfların varlığında doğruluk oranı kıyaslanması

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı	Parametreler
					c / gamma/ kernel
Bonn	ADD (4. Seviye, db2)	AxBxC	DVM	0,9041	1/ 0,1 / rbf
		AxBxD		0,891	10 /0,1 /rbf
		AxBxE		0,9166	10 /0,1 /rbf
		AxDxE		0,92	50 /0,1 /rbf
		AxCxE		0,929	50/ 0,1 /rbf
		BxDxE		0,9458	50/ 0,1 /rbf
		BxCxE		0,95	10/ 0,1 /rbf

Tablo 4.19’da gösterildiği üzere farklı sınıfların oluşturduğu kümelerin doğruluk oranları diğer benzer sınıfların oluşturduğu kümelere göre yüksek çıkmıştır. Bunun yanında benzer AxBxC ile AxBxE’nin doğruluk oranları karşılaştırıldığında E sınıfını bulduran kümenin daha yüksek doğruluk oranı ile sınıflandırıldığı anlaşılmaktadır.

Üçlü oluşturulan kümelerin yanı sıra Literatürde de özellikle tercih edilen farklı kümelemelerle birlikte de sonuçlar denenmiştir. Denenen kümelemeler doğruluk oranlarıyla birlikte Tablo 4.20’ye eklenmiştir. AxBxCxDxE kümesi için beş sınıf bulunmaktadır. ABCDxE kümesi içinse ABCD bir sınıf etiketiyle E ise ikinci sınıf etiketliyle eklenmişti. Böylelikle 2 sınıf oluşmuştur. ABxCDxE kümesinde ise AB bir sınıfı, CD ikinci sınıfı E ise üçüncü sınıfı temsil edecek şekilde etiketlenmiş ve sınıflandırma doğruluk oranı ölçülmüştür.

**Tablo 4.20:** Farklı sınıflar arasında DVM ile sınıflandırma doğruluk oranlarının ölçülmesi.

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı	Parametreler
					c / gamma/ kernel
Bonn	ADD (4. Seviye, db2)	AxBxCxDxE	DVM	0,7975	10/ 0,125/ rbf
		ABCDxE		0,9675	1/0,125/ rbf
		ABxCDxE		0,9325	50/0,1/ rbf

Tablo 4.20’den de anlaşıldığı üzere en yüksek sınıflandırma doğruluk oranını ABCDxE kümesi vermiştir. Sebebi ise bütün verilere karşılık nöbet halindeki hastanın verileri ile kıyaslanması olmuştur. Buna karşılık en düşük doğruluk oranı ise

AxBxCxDxE kümesinde elde edilmiştir. Sebebi ise birbirine benzeyen ikişer sınıfa sahip olması ve 5 kümeyi sınıflandırmak zorunda olmasıdır.

#### 4.1.2.2. DEÖS Kullanılarak Öznitelik Oluşturma

Bu kısımda ADD ile oluşturulan öznitelik kümeleri arasında en yüksek doğruluk oranına sahip öznitelik alt kümesi bulunmaya çalışılmıştır. Tablo 4.21’de DEÖS ile seçilen 10 özniteliğe sahip alt kümenin DVM ile elde edilen doğruluk oranı gösterilmiştir.

**Tablo 4.21 :** 10 özniteliğe sahip alt kümenin DVM ile sınıflandırma doğruluk oranı gösterimi

CR	İterasyon sayısı	Sınıflandırıcı	Doğruluk Oranı	Parametreler
				c/ gamma/ kernel
0,3	10000	DVM	0,8	10/ 0,2/ rbf
0,5	10000		0,7925	100/ 0,17/ rbf
0,7	10000		0,7875	500/ 0,2/ rbf
0,3	1000		0,82	100/ 0,15/ rbf
0,5	1000		0,79	500/ 0,1/ rbf
0,7	1000		0,8225	50/ 0,17/ rbf
0,3	100		0,8225	50/ 0,1/ rbf
0,5	100		0,8175	500/ 0,15/ rbf
0,7	100		0,815	50/ 0,1/ rbf

Bu tablo sonucunda en yüksek doğruluk oranının 0,7 ve 0,3 çaprazlama oranı ve farklı iterasyon sayıları ile bulunduğu gözlenmiştir. Bunun sebebi 24 öznitelik içerisinde seçilebilecek 10’lu alt küme sayılarının çok olmasıdır. Tablo 4.22’de 24 öznitelikten doğruluk oranını arttırmayı amaçlayan 15 elemanlı alt kümeler seçilmesi gösterilmiştir.



**Tablo 4.22:** 15 özniteliğe sahip alt kümenin DVM ile doğruluk oranı gösterimi

CR	İterasyon sayısı	Sınıflandırıcı	Doğruluk Oranı	Optimum parametreler
				c / gamma/ kernel
0,3	10000	DVM	0,8	10 /0,2/ rbf
0,5	10000		0,7925	100/ 0,17/ rbf
0,7	10000		0,7875	500/ 0,2/ rbf
0,3	1000		0,82	100/ 0,15/ rbf
0,5	1000		0,79	500/ 0,1/ rbf
0,7	1000		0,8225	50/ 0,17 /rbf
0,3	100		0,8225	50/ 0,1/ rbf
0,5	100		0,8175	500/0,15 /rbf
0,7	100		0,815	50/ 0,1/ rbf

Tablo 4.22'nin sonuçlarının Tablo 4.21'den farklı olmadığı gözlenmiştir. LDA ile sınıflandırmada bu durum oluşmamışken DVM'de böyle bir sonuç çıkmıştır. Burdan DVM için oluşturulmaya çalışılan alt kümenin sonuca büyük bir etkisi olmadığı DVM'nin daha iyisini bulmaya çalıştığı sonucu çıkarılmıştır. Tablo 4.23'de 24 öznitelikten sınıflandırma doğruluk oranını arttırmayı amaçlayan 20 elemanlı alt kümeler seçilmesi gösterilmiştir.

**Tablo 4.23:** 20 özniteliğe sahip alt kümenin DVM ile doğruluk oranı gösterimi

CR	İterasyon sayısı	Sınıflandırıcı	Doğruluk Oranı	Optimum parametreler
				c / gamma/ kernel
0,3	10000	DVM	0,8075	50/ 0,1/ rbf
0,5	10000		0,765	50/ 0,1/ rbf
0,7	10000		0,7875	50/ 0,1/ rbf
0,3	1000		0,78	10/ 0,125/ rbf
0,5	1000		0,7925	10/ 0,15/ rbf
0,7	1000		0,78	100/ 0,125/ rbf
0,3	100		0,8075	50/ 0,1/ rbf
0,5	100		0,795	10/ 0,125/ rbf
0,7	100		0,78	100/ 0,1/ rbf

Tablo 4.23'de gösterildiği gibi sınıflandırma doğruluk oranı diğer alt kümelerin doğruluk oranını geçecek şekilde çıkamamıştır. Aynı durum 24 öznitelikli kümeden 20 öznitelikli alt küme oluşturup doğruluk oranını LDA ile test edince de gözlenmiştir. Sebebi de 24 öznitelik arasından 20 elemana sahip sınırlı sayıda alt küme çıkarmasıdır.

#### 4.1.2.3. 1D-YİÖ Kullanılarak Öznitelik Oluşturma

Filtrelenen sinyalden 1D-YİÖ ile öznitelik kümesi oluşturulması ile test edilmiştir. Deneyler tıpkı ADD'deki gibi ikili, üçlü ve dörtlü sınıfların varlığında sınıflandırma doğruluk oranlarının karşılaştırılması ile oluşmuştur. İkili olarak A sınıfı ile diğer sınıfların varlığında DVM ile doğruluk oranı kıyaslanması Tablo 4.24'de gösterilmiştir.

**Tablo 4.24:** A sınıfı ile diğer sınıfların ikili oluşturdukları kümenin DVM ile doğruluk oranı

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırma	Doğruluk Oranı	Parametreler
					c/ gamma/kernel
Bonn	1D-YİÖ	AxB	DVM	0,86	10/ 0,1/ rbf
		AxC	DVM	0,8875	10/ 0,1/ rbf
		AxD	DVM	0,825	10/0,15 /rbf
		AxE	DVM	0,8	1/0,1 /rbf

Bu tabloya göre sınıflandırma sonucu LDA ile yapılan testlerin doğruluk oranına yakın çıkmıştır. Doğruluk oranları kıyaslaması B sınıfının ikili olarak diğer sınıflar ile kombinlenmesi ile elde edilen kümelerde denenmiştir. Sonuçlar Tablo 4.25'de gösterilmiştir.

**Tablo 4.25:** B sınıfı ile diğer sınıfların ikili oluşturdukları kümenin DVM ile doğruluk oranları

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kullanılan Küme	Sınıflandırma	Doğruluk Oranı	Parametreler
					c/ gamma/ kernel
Bonn	1D-YİÖ	BxC	DVM	0,9125	10/ 0,1/ rbf
		BxD		0,875	10/ 0,1/ rbf
		BxE		0,8125	1/ 0,1/ rbf

Tablo 4.25'de en yüksek sınıflandırma doğruluk oranı Tablo 4.24'deki gibi C sınıfıyla yapılan kümelendirme sonucu bulunmuştur. Sonraki en yüksek doğruluk oranı C sınıfıyla oldukça benzerliğe sahip olan D sınıfı ile olan kümeden elde edilmiştir. İkili olarak C ve D sınıflarının ayrı ayrı diğer sınıflar ile oluşturdukları kümelerin doğruluk oranları kıyaslanması Tablo 4.26'da gösterilmiştir.

**Tablo 4.26:** C ve D sınıflarının diğer sınıflar varlığında ayrı ayrı ikili kombinler ile oluşturdukları kümelerin DVM ile doğruluk oranları

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırıcı	Doğruluk Oranı	Parametreler
					c/ gamma/ kernel
Bonn	1D-YİÖ	CxD	DVM	0,58125	10/0,1/ rbf
		CxE		0,71125	1/ 0,1/ rbf
		DxE		0,7375	10/ 0,1/ rbf

Tablo 4.26’da da gösterildiği üzere en düşük doğruluk oranı birbirleriyle çok ortak özelliği olan CxD kümeleridir. Diğer kümelermeler CxD kümesinin doğruluk oranının üstünde çıkmıştır. Üçlü kombinler halinde oluşturulan kümelerin doğruluk oranlarının ölçülmesi Tablo 4.27’de gösterilmiştir.

**Tablo 4.27:** Üçlü kombinlerin DVM ile doğruluk oranlarının kıyaslanması

Kullanılan Veri	Öznitelik Seçim Yöntemi	Kümeler	Sınıflandırma	Doğruluk Oranı	Parametreler
					c/ gamma/ kernel
Bonn	1D-YİÖ	AxBxC	DVM	0,808	1/ 0,1/ rbf
		AxBxD		0,7791	1/ 0,1/ rbf
		AxBxE		0,745	10/ 0,1/ rbf
		AxCxD		0,591	10/ 0,1/ rbf
		AxCxE		0,6875	1/ 0,1/ rbf
		AxDxE		0,7125	10/ 0,1/ rbf
		BxCxD		0,5875	1/ 0,1/ rbf
		BxCxE		0,675	10/ 0,1/ rbf
		BxDxE		0,65	10/ 0,1/ rbf
		CxDxE		0,5125	10/ 0,1/ rbf

Tablo 4.27’deki gösterildiği üzere üçlü kümelere en düşük sonucu veren küme üçü de hasta sınıfa sahip olan CxDxE kümesidir. İkili ve üçlü kombinlerin ardından dörtlü kombinler oluşturulup sınıflandırma doğruluk oranları kıyaslanmıştır.

#### 4.1.3. Bütün Yöntemlerin Karşılaştırılması

Bonn verisi sonucu farklı veri kümeleri elde edilmiş ve hepsinin sınıflandırma doğruluk oranları kıyaslanmıştır. AxE ve BxE için ADD ve 1D-YİÖ karşılaştırılması Tablo 4.28'deki gibidir.

**Tablo 4.28:** AxE ve BxE için ADD ve 1D-YİÖ doğruluk oranları karşılaştırılması

Kullanılan Veri	Öznitelik Seçimi	Sınıflandırma	Kullanılan Küme	Doğruluk Oranı
Bonn	ADD	LDA	AxE	0,9799
		DVM		0,9875
	1D-YİÖ	LDA	AxE	0,885
		DVM		0,8
		LDA	BxE	0,795
		DVM		0,8125
	ADD	LDA	BxE	0,98
		DVM		0,9812

Tablo 4.28'e göre AxE kümesinin en yüksek doğruluk oranı veren işlem ADD ve DVM'nin birlikte kullanılması sonucu oluşmuştur. BxE kümesi için de ADD ve DVM'nin birlikte kullanılması ile bulunmuştur. Bütün sınıfları tek bir tabloda kıyaslamak için A, B, C, D ve E sınıfları ele alınıp tek bir tabloda kıyaslama işlemi yapılmıştır.

**Tablo 4.29:** Beşli küme için kullanılan bütün yöntemlerin kıyaslanması.

Kullanılan Veri	Öznitelik Seçimi	Sınıflandırma	Kullanılan Küme	Doğruluk Oranı
Bonn	ADD	LDA	AxBxCxDxE	0,722
		DVM		0,7975
	1D-YİÖ	LDA		0,57
		DVM		0,555
	ADD +DEÖS	LDA		0,74
		DVM		0,8225
	ADD + TBA	LDA		0,7669

Tablo 4.29'da elde edilen sonuçlara göre beşli kümenin en yüksek doğruluk oranı verdiği yöntem ADD +DEÖS ile DVM'nin kullanılması sonucu oluşmuştur. DEÖS'ün doğruluk oranını arttırdığı görülmektedir.

## 4.2. BCI Verisi İçin Bulgular

BCI yarışma verisinde ise 3 denek bulunmaktadır; k3b, k6b ve l1b. Bu tezde k3b deneğinin verisi üzerinde çalışılmıştır. K3b deneği 4 sınıflı bir veridir. Denek ekranda gösterilen işaretlere göre zihninde sol el, sağ el, ayak ve dilini oynattığını düşünmektedir. Toplamda 4 sınıfa sahip bir veri setidir.

BCI yarışma veri kümesine bant geçiren filtresi uygulanmıştır. Bant geçiren filtresi yüksek ve alçak bant frekansları arasının dışında kalan alanları bastıran ve sadece o arada kalan verileri ortaya çıkaran bir yöntemdir. 8 Hz'den küçük frekanslı sinyalleri ve 30 Hz'den büyük frekanstaki sinyalleri bastırmıştır. Örnekleme frekansı 250 Hz olan veri ile 4 saniye boyunca yapılan deney süresi hesaplanmış ve 1000 veri noktasının bir deney süresi olduğu anlaşılmıştır. Örnekte toplam 360 deneme olduğu için her bir denemenin 1000x60 boyutunda olduğu anlaşılmıştır. Veri etiketli hale dönüştürülüp 4 sınıfa ayrılmıştır. Her bir sınıf için 90000x60'lık bir matris oluşturulmuştur.

### 4.2.1. CSP Uygulanması

CSP'nin ikili sınıflandırma kuralına göre ikili kombinler halinde CSP işlemi sonucu uzamsal desen bulunmuştur. Bulunan matrisin tersi alınarak sütunları incelenmiş ve sütunlarda mutlak maksimum değeri veren elemanın satır sayısı aktif kanalı belirlemiştir. CSP ikili sınıflarda uygulanınca 6 adet uzamsal örüntü matrisi oluşmuştur (1x2, 1x3, 1x4, 2x3, 2x4, 3x4). Her birinden en yüksek değerli 10 kanal seçilmiş ve ortak olanlardan 12 tanesi aktif kanal olarak alınmıştır. Uygulanan yöntemler sonucu özet tablo bölümün sonunda gösterilmiştir.

### 4.2.2. ADD Kullanılarak Öznitelik Oluşturma

Seçilen 12 kanala beşinci dereceden ADD uygulanmış ve cD3, cD4, cD5, cA5 alt bant katsayıları alınmış ve her bir alt bant katsayısından istatistiksel özellikler çıkarılmıştır. Bu özellikler; ortalama, minimum, maksimum, enerji, eğrilik (skewness), basıklık (kurtosis), standart sapma, orta değer, mutlak toplam ve varyanstır. Toplamda 432 öznitelik çıkarılmıştır (9 öznitelik özelliği x 4 alt bant x 12 kanal). Öznitelikler kullanılarak sınıflandırma işlemi yapılmıştır.

#### 4.2.3. DEÖS Kullanılarak Öznitelik Oluşturma

ADD ile öznitelik oluşturulmuş kümeden Diferansiyel evrim algoritması ile de en iyi öznitelik alt kümesi seçilmeye çalışılmış ve sınıflandırma DEÖS algoritmasının farklı parametreleri ile denenmiştir.

#### 4.2.4. Ki Kare ile Öznitelik Oluşturma

ADD sonucu elde edilen matris 0-1 arasında normalize edilerek ki-kare testi uygulanmış ve sınıflandırmaya sokulmuştur.

#### 4.2.5. 1D-YİÖ İle Öznitelik Oluşturma

ADD metoduna ek olarak 1D-YİÖ de kullanılmıştır. Bu yöntem her bir kanaldan 256 özellik çıkararak oluşmuştur. 60 kanal için  $256 \times 60 = 15360$  öznitelik elde edilmiştir. Matrisin boyutu  $360 \times 15360$  olmuş ve sınıflandırmaya sokulmuştur. Veriye 1D-YİÖ kullanılarak her bir örneğin 9'lu olarak ayrılması sağlanmıştır. 9'lu kümelerin ortalarındaki elemanla kıyaslanarak 1 ve 0 değerleri elde edilmiştir. Eğer komşu eleman merkez elemandan büyükse 1 değilse 0 değerini almıştır. Bu sayede 256 öznitelik elde edilmiştir. Bu öznitelikler kullanıcıya hangi komşunun merkezi pikselden daha büyük olduğu bilgisini vermiştir. Bu sayede BCI verisinin tek sınıfı için  $180 \times 15360$  boyutunda küme elde edilmiştir. Sınıf etiketi eklenince  $180 \times 15361$  boyutlu matris oluşmuştur.

BCI Yarışması veri seti için sırasıyla denenilen işlemlerin özeti şöyledir;

- Filtreleme -> CSP -> ADD -> LDA
- Filtreleme -> CSP -> ADD ->Ki Kare ->LDA
- Filtreleme -> CSP -> ADD ->DEÖS ->LDA
- Filtreleme -> 1D -YİÖ -> LDA

#### 4.2.6. Bütün Yöntemlerin Karşılaştırılması

Bu tez sonucu elde ettiğimiz sonuçlara göre Bonn verisinin BCI veri setinden daha düzenli ve doğru veriler içerdiği söylenebilmiştir. Kullandığımız CSP ve 1D-YİÖ, ADD, DEÖS gibi literatürde güçlü olduğu bilinen yöntemlere Bonn verisinin iyi sonuç verdiği fakat BCI verisinin vermediği görülmüştür. Bu tezde BCI verisi içinde farklı öznitelik seçim metotları denenmiştir. Sonuçlar Tablo 4.30'da gösterildiği şekilde çok yüksek çıkmamıştır.

**Tablo 4.30:** BCI verisi için denenen metotlar ve doğruluk oranları

Kullanılan Veri	Öznitelik Seçim Yöntemi	Sınıflandırma	Kullanılan Küme	Doğruluk Oranı
BCI Yarışma 3 Veri set 3a	CSP +ADD (5. Seviye, db4)	LDA	K3b (4 sınıf)	0,26
	CSP+ADD (5. Seviye, db4) +Ki Kare			0,30
	1D-YİÖ			0,28

BCI verisi için denenen Diferansiyel Evrim Algoritması ile iterasyon sayısı 10 000 iken CR ise 0.8 - 0.3 iken elde edilen sonuçlar Tablo 4.31 'deki gibidir.

**Tablo 4.31:** BCI verisi DEÖS 'ün farklı CR değerlerine göre doğruluk oranı kıyaslaması

CR	İterasyon sayısı	Sınıflandırıcı	Doğruluk Oranı
0,8	10000	LDA	0,26
0,5			0,31
0,3			0,28

Tablo 4.31'de gösterildiği üzere DEÖS kullanımının sınıflandırma doğruluk oranını bir miktar arttırdığı fakat yeterli seviyeye gelmediği gözlenmiştir.

## 5. SONUÇLAR VE ÖNERİLER

Literatürde en çok kullanılan EEG veri setlerinden olan Bonn ve BCI veri kümeleri kullanılarak öznitelik çıkarım ve seçim metodlarının başarıya etkileri gözlenmiştir. Kullanılan Evrimsel Öznitelik Seçim Metodu ile çıkarılan özniteliklerden alt öznitelik kümesi seçilerek DVM ve LDA ile sınıflandırma yapılmıştır. Kullanılan sınıflandırıcıların ürettiği doğruluk değerlerine göre seçilen öznitelik çıkarım metodları arasında performans karşılaştırılması yapılabilmektedir.

İlk olarak Bonn verisinin sonuçlarına baktığımızda Bonn verisi işlenmesi oldukça kolay bir veri seti olup doğruluk oranı hem literatürde hem de tez sonuçlarımızda yüksek çıkmıştır. Bunun sebeplerinden birisi de veri setini oluşturan uzmanlar tarafından manuel olarak bozulma tespiti (artifact selection) yapılmış olmasıdır. Bonn verisinde daha az örnek bulunması sebebiyle DVM'nin daha başarılı sonuç verdiği tespit edilmiştir. Bonn verisine uygulanan yöntemlerden biri olan ADD'nin sinyal verileri için etkili bir öznitelik çıkarım yöntemi olduğu tespit edilmiştir. İkili kümelerin sınıflandırılmasında literatürdeki gibi yüksek sonuçlar elde edilmiştir. Bu çalışmada bulunan A ve E sınıflarının birleşimiyle oluşan kümenin doğruluk oranı 0,9875 iken literatürde %98, %99, %100, %91 ve %96 gibi değerler bulunmuştur. Bu durumda bu çalışma ikili kümelere bazı sonuçları geçmiş bazılarında ise çok yaklaşmıştır. ADD ile üçlü kümelerin varlığında yapılan çalışmada ise bilinen birçok çalışmayı geçmiştir. Literatürde AxCxE ile yapılan deneyde %80 sonucuna karşılık bu çalışmada %92, BxCxE ile yapılan deneyde literatürdeki sonuç %78,7 iken bu tezde %95 olmuştur. Literatürde %80 ve %98 olarak bulunan AxDxE kümesinin sonucu bu çalışmada %92 olarak bulunmuştur. Ayrıca %79,3 olan BxDxE kümesinin doğruluk oranı bu çalışmada %94 olarak bulunmuştur. Ayrıca ABCDxE gibi nöbet halindeki ve diğer denekleri kıyaslayan kümede bilinen sonuç %91 iken, bu çalışmada %96 olarak bulunmuştur. ADD ile elde edilen 5 sınıflı kümeye DEÖS uygulanarak LDA ile %72 olan çaprazlama oranının 0,7 ve iterasyon sayısının 10000 olduğu deney ortamında doğruluk oranı %74'e çıkmıştır. DVM ile %79 olan doğruluk oranı çaprazlama oranının 0,7 iterasyon sayısının 1000 olduğu deneyde %82'ye ulaştığı gözlenmiştir. En yüksek doğruluk oranını veren DEÖS yöntemi bu oranı 15 elemanlı alt kümeye sahip olduğunda elde etmiştir. DEÖS ile öznitelik seçim yönteminde 24 elemanlı



kümeden 10, 15 ve 20 elemanlı alt kümeler seçilerek kıyaslanan sınıflandırma doğruluk oranları 20 elemanlı kümede çok yüksek sonuçlar vermemiştir. Bunun sebebi 24 elemanlı kümeden sınırlı sayıda 20 elemanlı alt küme oluşmasıdır. DEÖS ile seçilen 10 ve 15 elemanlı alt kümelerin sınıflandırma doğruluk oranlarının DVM ile ölçümünde değişim gözlenmemiştir. 10 elemanlı alt kümede de 15 elemanlı alt kümede de maksimum kabul ettiğimiz %82'ye ulaşmıştır. Bu durumun sebebi DVM'nin var olan en iyi sonucu bulmaya LDA'dan daha elverişli olmasıdır. Bu bakımdan DVM bu problem için LDA'dan daha iyidir denilebilmiştir. Ayrıca DEÖS değişkenlerinden biri olan iterasyon sayısının düşük olması doğru sonucun bulunmamasına neden olmaktadır. 100 ile yapılan deneyler 1000 ve 10000 ile yapılan deneylere göre daha düşük sınıflandırma doğruluk oranı veren alt kümeler oluşturmuştur. TBA ile boyut azaltma işlemi uygulandığında ise bileşen sayısı 10'a düşürüldükten sonra LDA ile doğruluk oranı  $0,7669 \pm 0,1525$  olarak bulunmuştur. ADD sonucu elde edilen veri kümesine TBA uygulandığı için doğruluk oranını arttırmıştır. Çünkü ADD ile elde edilen öznitelikler istatistiksel verilerdir ve TBA uygulanması sonucu doğruluk oranının artması birbirleriyle bağlantılı olduklarını göstermektedir. Boyut azaltma ile sınıflandırma doğruluk oranı arttırılmıştır. 1D-YİÖ yöntemi bu çalışmada ADD'ye bir alternatif olarak düşünülmüştür. Literatürde bulunan sonuçlara benzer sonuçlar elde edilmiştir. İkili ve üçlü kombinlerde ADD ile yarışabilecek sonuçlar vermesine rağmen beşli kümelerde sınıflandırma doğruluk oranı açısından ADD'den daha düşük sonuçlar vermiştir. Ayrıca sınıflandırıcıya verilen kümenin boyutunun ADD sonucu oluşan kümenin boyutundan büyük olduğu gözlenmiştir. Bu sebeple sınıflandırma işlemi süresi daha büyük verilerde daha uzun olacağı tahmin edilmektedir. Literatürde bu yöntemle beşli küme için sınıflandırma doğruluk oranı bulunmamaktadır. Bu eksikliğin giderilmesi için bu yöntemin doğruluk oranının arttırılması ileriki çalışmalar için düşünülmüştür.

Son olarak BCI verisi sonuçları incelendiğinde bir yarışma verisi olması ve yarışmacıları özellikle zorlamak amacıyla oluşturulduğu düşünülünce işlenmesi ve sınıflandırılması zor bir veri olduğu tahmin edilmiştir. Literatürde iyi sonuç verdiği bilinen kanal seçme ve öznitelik çıkarım yöntemleri uygulanmasına rağmen yüksek sonuçlar elde edilememiştir. Bu sebeple BCI verisinde bozulma olduğu düşünülmektedir. Bu sorunun çözümü olarak ileriki çalışmalarda bozulma tespiti yapılması düşünülmüştür.

## KAYNAKLAR

1. EEG Time Series. [Çevrimiçi] 2005. [Alıntı Tarihi: 7 10 2018.] [http://epileptologie-bonn.de/cms/front\\_content.php?idcat=193&lang=3&changelang=3](http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3).
2. BCI Competition 3. BCI Competition 3. [Çevrimiçi] 12 Aralık 2004. [Alıntı Tarihi: 07 10 2018.] <http://www.bbc.de/competition/iii/>.
3. Tohru, Okanishi. S18-1. Progress in EEG analysis for epilepsy and the outlook for the future. Clinical Neurophysiology. 5, May 2018, Cilt 129, 5.
4. DoktorX. DoktorX Sağlık. [Çevrimiçi] 22 11 2017. [Alıntı Tarihi: 14 12 2018.] <https://www.doktorix.com/eeg-elektroensefalografi-nedir-neden-ve-ne-zaman-cekilir-eeg-yorumlama/>.
5. Teplan, Michal. Fundamentals of EEG measurement. Measurement science review. 2, 2002, Cilt 2, 1-11.
6. Ö.Demir. EEG Dalgalarının Wavelet (Dalgacık) Dönüşümü ile Değerlendirilmesi. Yayınlanmamış Yüksek Lisans Tezi. Dumlupınar Üniversitesi Fen Bilimleri Enstitüsü Dergisi. 2008.
7. Daşdağ, S. Elektroensefalografinin (EEG) Biyofizik. Dicle Üniversitesi. [Çevrimiçi] 07 01 2017. [Alıntı Tarihi: 08 12 2018.] <http://www.dicle.edu.tr/a/dasdag/EEG.ppt>.
8. R., Q. Quiroga. Quantitative Analysis of EEG Signals: Time Frequency Methods and Chaos Theory. Yayınlanmamış Doktora Tezi. Verlag nicht ermittelbar. 1998.
9. Sezer, Esmâ. Epilepsi Teşhisi için EEG Sinyal Analizi. Yayınlanmamış Yüksek Lisans Tezi. Selçuk Üniversitesi Fen Bilimleri. 2008.
10. Md, Mamun Or Rashid ve Ahmad, Mohiuddin. Epileptic Seizure Classification using Statistical Features of EEG Signal. In: 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE), IEEE. 2017, Cilt 308-312.
11. Ahmadi, Amirmasoud ve Vahid Shalchyan Reza, Daliri Mohammad. A New Method for Epileptic Seizure Classification in EEG Using Adapted Wavelet Packets. Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT),IEEE. 1-4, 2017.
12. Ren, Weijie. Efficient feature extraction framework for EEG signals classification. Intelligent Control and Information Processing (ICICIP), Seventh International Conference on. IEEE, 2016. 2016, 167-172.
13. Ramakrishnan, S. ve As, Muthanatha Murugavel. Epileptic seizure detection using fuzzy-rules-based sub-band specific features and layered multi-class SVM. Pattern Analysis and Applications. 1, 2018, Cilt 16.
14. Physio Bank. [Çevrimiçi] [Alıntı Tarihi: 12 11 2018.] (<http://physionet.org/cgi-bin/atm/ATM>).
15. Salem, Osman, Naseem, Amal ve Mehaoua, Ahmed. Epileptic Seizure Detection From EEG Signal using Discrete Wavelet Transform and Ant Colony classifier.

- Communications (ICC), 2014 IEEE International Conference on. IEEE, 2014, 3529-3534.
16. Bose, Suvadeep, Rama, V. ve Ch, Rama Rao. EEG signal analysis for Seizure detection using Discrete Wavelet Transform and Random Forest. Computer and Applications (ICCA), 2017 International Conference on. IEEE. 2017, 369-378.
  17. Duo Chen, Suren Wan , Jing Xiang, Forrest Sheng Bao. A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. Plos one. 12, 2017, Cilt 12, 3.
  18. Yılmaz Kaya, Murat Uyar, Ramazan Tekin, Selçuk Yıldırım. 1D-local binary pattern based feature extraction for classification of epileptic EEG signals. Applied Mathematics and Computation. 2014, Cilt 243, 209-219.
  19. Acharya U Rajendra ve Sree, S Vinitha and Alvin, Ang Peng Chuan and Suri, Jasjit S. Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. Expert Systems with Applications. 10, 2012, Cilt 39, 9072-9078.
  20. Ahammad, Nabeel, Fathima, Thasneem ve Joseph, Paul. Detection of epileptic seizure event and onset using EEG. BioMed research international 2014. 2, 2014, Cilt 2014, 7.
  21. Kavita Mahajan, M. R. Vargantwar, Sangita M. Rajput. Classification of EEG using PCA, ICA and Neural Network. International Journal of Engineering and Advanced Technology. 1, 2011, Cilt 1, 80-83.
  22. Abdulhamit Subasi, Gursoy M. Ismail. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert systems with applications. 37, 2010, Cilt 12, 8659-8666.
  23. Muhammad Zeeshan Baig, Nauman Aslam, Hubert P.H.Shum, Li Zhang. Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG. Expert Systems with Applications. 30 December 2017, Cilt 90, 184-195.
  24. BCI Competition III - Berlin BCI. BCI Competition 3 4a. [Çevrimiçi] 12 12 2004. [Alıntı Tarihi: 12 11 2018.] [http://www.bbc.de/competition/iii/#data\\_set\\_iva](http://www.bbc.de/competition/iii/#data_set_iva).
  25. Saka Kübra, Önder Aydemir, Mehmet Öztürk. Classification of EEG signals recorded during right/left hand movement imagery using Fast Walsh Hadamard Transform based features. telecommunications and Signal Processing (TSP), 2016 39th International Conference on. IEEE. 2016, 413-416.
  26. BCI 2.Competition 3. [Çevrimiçi] 1 Mayıs 2003. [Alıntı Tarihi: 12 11 2018.] <http://webcache.googleusercontent.com/search?q=cache:http://www.bbc.de/competition/ii/>.
  27. Yijun Wang, Shangkai Gao, Xiaonog Gao. Common Spatial Pattern Method for Channel Selection in Motor Imagery Based Brain-computer Interface. Engineering in medicine and biology society, 2005. IEEE-EMBS 2005. 27th Annual international conference of the. IEEE, 2006, 5392-5395.

28. Koprinska, Irena. Feature Selection for Brain-Computer Interfaces. Pacific-Asia conference on knowledge discovery and data mining. Springer, Berlin, Heidelberg, 2009, 106-117.
29. Indah, Agustien Siradjuddin. Feature Selection with Genetic Algorithm for Alcoholic detection using Electroencephalogram. Sustainable Information Engineering and Technology (SIET), 2017 International Conference on. IEEE, 2017, 230-234.
30. UCI Machine Learning Repository. UCI Machine Learning Repository EEG. [Çevrimiçi] [Alıntı Tarihi: 13 11 2018.] <https://archive.ics.uci.edu/ml/datasets/eeg+database>.
31. CHB-MIT Scalp EEG Database. physionet.org. [Çevrimiçi] [Alıntı Tarihi: 08 12 2018.] <https://www.physionet.org/pn6/chbmit/>.
32. Hamad, Asmaa. Feature extraction of epilepsy EEG using discrete wavelet transform. 2016 12th international computer engineering conference (ICENCO). IEEE, 2016, 190-195.
33. Kai Keng Ang, Zheng Yang Chin, Chuanchu Wang, Cuntai Guan, and Haihong Zhang. Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b. *Frontiers in neuroscience*. 6, 2012, Cilt 39.
34. Dornhege, Guido. Increase information transfer rates in BCI by CSP extension to multi-class. *Advances in neural information processing systems*. 2004, 733-740.
35. Siuly Siuly, Li Yan, Zhang Yanchun. EEG Signal Analysis and Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 11, Techniques and Applications, Cilt 2, 141-144.
36. Abdulhamit, Subasi. EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*. 32, 2007, Cilt 4, 1084-1093.
37. Ojala Timo, Matti Pietikainen, Topi Maenpaa. Gray scale and rotation invariant texture classification with local binary patterns. *European Conference on Computer Vision*. Springer, Berlin, Heidelberg. 2000, 404-420.
38. Kurt Burçin, Nabiye V. Vasif. Down syndrome recognition using local binary patterns and statistical evaluation of the system. *Expert Systems with Applications*. 38, 2011, Cilt 7, 8690-8695.
39. Navin Chatlani, J. Sorahgan John. Local binary patterns for 1-D signal processing. *Signal Processing Conference, 2010 18th European*. IEEE, 2010, 95-99.
40. Pearson, F.R.S. Karl. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. 2, 1901, Cilt 11, 559-572.
41. N.Aronszajn. Theory of Reproducing Kernels. *Transactions of the American mathematical society*. 68, 1950, Cilt 3, 337-404.
42. SVM. SVM. [Çevrimiçi] [Alıntı Tarihi: 04 12 2018.] <http://www.svms.org/history.html>.

43. E. Boser Bernhard, M. Guyon Isabelle, N. Vapnik Vladimir. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory. ACM, 1992, 144-152.
44. Cortes Corinna, Vapnik Vladimir. Support-vector networks. Machine Learning. 3, 1995, Cilt 20, 273–297.
45. Duman, Fırat. ‘Destek Vektör Makineleri Yardımıyla İnvaziv EEG Verilerinden Epileptik Nöbet Tahmini. Mustafa Kemal Üniversitesi Fen Bilimleri Enstitüsü. 2013.
46. Kemp, Freda. Applied multiple regression/correlation analysis for the behavioral sciences. Journal of the Royal Statistical Society. Series D (The Statistician), 2003, Cilt 52, 4.
47. Kenneth Price, M. Storn Rainer, A. Lampinen Jouni. Differential Evolution A Practical Approach to Global Optimization. Springer Science & Business Media, 2006.
48. Rami N.Khushaba, Ahmed Al-Ani, Adel Al-Jumaily. Feature subset selection using differential evolution and a statistical repair mechanism. Expert Systems with Applications. 9, 2011, Cilt 38, 11515-11526.