

**KARADENİZ TEKNİK ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**KİSMİ EN KÜÇÜK KARELER REGRESYON YÖNTEMİYLE UYKU EEG'SİNİN**  
**SINIFLANDIRILMASI**

**YÜKSEK LİSANS TEZİ**

**Elektrik-Elektronik Mühendisi Kübra EROĞLU**

**OCAK 2013**  
**TRABZON**

**KARADENİZ TEKNİK ÜNİVERSİTESİ**

**FEN BİLİMLERİ ENSTİTÜSÜ**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**KİSMİ EN KÜÇÜK KARELER REGRESYON YÖNTEMİYLE UYKU EEG'SİNİN  
SINIFLANDIRILMASI**

**Elektrik-Elektronik Mühendisi Kübra EROĞLU**

**Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünde  
"ELEKTRİK YÜKSEK MÜHENDİSİ"  
Unvanı Verilmesi İçin Kabul Edilen Tezdir.**

**Tezin Enstitüye Verildiği Tarih : 18.12.2012  
Tezin Savunma Tarihi : 03.01.2013**

**Tez Danışmanı : Prof. Dr. Temel KAYIKÇIOĞLU**

**Trabzon 2013**

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü  
Elektrik Elektronik Mühendisliği Anabilim Dalında  
Kübra EROĞLU tarafından hazırlanan

**KİSMİ EN KÜÇÜK KARELER REGRESYON YÖNTEMİYLE UYKU EEG'SİNİN  
SINIFLANDIRILMASI**


başlıklı bu çalışma, Enstitü Yönetim Kurulunun 18/12/2012 gün ve 1486 sayılı  
kararıyla oluşturulan jüri tarafından yapılan sınavda  
**YÜKSEK LİSANS TEZİ**  
olarak kabul edilmiştir.

**Jüri Üyeleri**

**Başkan : Prof. Dr. Temel KAYIKÇIOĞLU**

  
.....

**Üye : Doç. Dr. Ali GANGAL**

  
.....

**Üye : Doç. Dr. Cemal KÖSE**

  
.....

**Prof. Dr. Sadettin KORKMAZ**

**Enstitü Müdürü**

## ÖNSÖZ

Bu tez çalışması, Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı'nda Yüksek Lisans Tezi olarak hazırlanmıştır. Uyku EEG işaretlerinin sınıflandırılma işleminin gerçekleştirildiği bu çalışmada uyku EEG sinyal analizi ile elde edilen verilerin Kısmi En Küçük Kareler Regresyon yöntemi ile sınıflandırılması ve sonuçların farklı iki sınıflandırma yöntemi ile karşılaştırılması gerçekleştirilmiştir.

Tez çalışma sürecimin her aşamasında bana yol gösterip bilimsel bilgi ve desteğini, esirgemeyen saygıdeğer hocam ve değerli danışmanım Prof. Dr. Temel KAYIKÇIOĞLU'na sonsuz teşekkürlerimi sunarım.

Çalışmam boyunca gösterdikleri sabır ve anlayış ile hep yanımda olup hayatım boyunca desteklerini hiç esirgemeyen başta annem ve babam olmak üzere bütün AİLEM'e tüm içtenliğimle teşekkür eder ve yapmış olduğum bu tez çalışmasının faydalı olmasını temenni ederim.

Kübra EROĞLU

Trabzon 2013

## TEZ BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “Kısmi En Küçük Kareler Regresyon Yöntemi ile Uyku EEG’sinin Sınıflandırılması” başlıklı bu çalışmayı baştan sona kadar danışmanım Prof. Dr. Temel KAYIKÇIOĞLU’nun sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim.  
03/01/2013

Kübra EROĞLU

## İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	III
TEZ BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET .....	VII
SUMMARY .....	VIII
ŞEKİLLER DİZİNİ .....	IX
TABLolar DİZİNİ.....	XI
SEMBOLLER DİZİNİ .....	XII
1. GENEL BİLGİLER .....	1
1.1. Giriş .....	1
1.2. EEG.....	3
1.2.1. EEG İşaretlerinin Ölçülmesi.....	4
1.2.2. EEG Elektrotlarının Bağlantı Şekilleri .....	5
1.2.3. EEG İşaretlerinin Kapsadığı Bantlar .....	8
1.3. Uyku.....	10
1.3.1. Laboratuvarlardada Uyku Ölçümü .....	11
1.3.2. Uyku EEG'si.....	12
1.3.3. Uyku Evreleri.....	13
1.4. Literatür Araştırması.....	18
2. YAPILAN ÇALIŞMALAR.....	22
2.1. Materyal .....	22
2.2. Yöntem.....	23
2.2.1. Ön İşleme .....	23
2.2.1.1. Normalizasyon .....	23
2.2.1.2. Filtreleme .....	24
2.2.2. Öz Nitelik Çıkarma.....	26
2.2.2.1. Zaman Serileri Analizi.....	27
2.2.2.2. Box-Jenkins Modelleri.....	28
2.2.2.2.1. Box-Jenkins Modelinin Kurulumu .....	28
2.2.2.2.2. Otokovaryans ve Otokorelasyon Fonksiyonları .....	29
2.2.2.2.3. Otoregresif (Auto Regressive-AR) Model.....	31

2.2.2.2.3.1.	Yule-Walker Yöntemi.....	31
2.2.2.2.3.2.	Burg Yöntemi .....	33
2.2.3.	Tahmin ve Etme ve Sınıflandırma .....	36
2.2.3.1.	Kısmi En Küçük Kareler Regresyon Yöntemi .....	36
2.2.3.1.1.	Çoklu Doğrusal Regresyon.....	37
2.2.3.1.1.1.	En Küçük Kareler Yöntemi .....	38
2.2.3.1.1.2.	Çoklu Bağlantı .....	39
2.2.3.1.2.	Ayarlama.....	40
2.2.3.1.2.1.	Veri Sıkıştırma.....	41
2.2.3.1.2.2.	Bilineer Yöntemler .....	42
2.2.3.1.3.	Kısmi En Küçük Kareler Yöntemi.....	43
2.2.3.1.4.	Kısmi En Küçük Kareler Regresyon Modeli.....	46
2.2.3.1.4.1.	KEKKR Model Geçerliliğinin İncelenmesi.....	49
2.2.3.1.5.	Kısmi En Küçük Kareler Regresyonunda Algoritmalar .....	50
2.2.3.1.5.1.	NIPALS Algoritması .....	50
2.2.3.1.5.2.	SIMPLS Algoritması .....	52
2.2.3.2.	k-En Yakın Komşuluk (k-NN) Yöntemi .....	53
2.2.3.3.	Bayes Yöntemi.....	55
3.	BULGULAR VE SONUÇLAR.....	59
3.1.	Öz Niteliklere Ait Bulgular.....	67
3.2.	Tahmin Etme ve Sınıflandırmaya Ait Bulgular .....	71
3.2.1.	Kısmi En Küçük Kareler Regresyon Yöntemine Ait Bulgular.....	71
3.2.2.	k-En Yakın Komşuluk Yöntemine Ait Bulgular .....	75
3.2.3.	Bayes Yöntemine Ait Bulgular.....	76
3.2.4.	Yöntemin Performansına İlişkin Bulgular .....	80
4.	TARTIŞMA VE ÖNERİLER.....	86
5.	KAYNAKLAR .....	87
ÖZGEÇMİŞ		

Yüksek Lisans Tezi

ÖZET

KİSMİ EN KÜÇÜK KARELER REGRESYON YÖNTEMİYLE UYKU EEG'SİNİN  
SINIFLANDIRILMASI

Kübra EROĞLU

Karadeniz Teknik Üniversitesi  
Fen Bilimleri Enstitüsü  
Elektrik-Elektronik Mühendisliği Anabilim Dalı  
Danışman: Prof. Dr. Temel KAYIKÇIOĞLU  
2013, 91 Sayfa

Bu çalışmada amaçlanan Elektroensefalografi (EEG) uyku kayıtlarında uyku ile uyanıklık evrelerinde gözlemlenen beyin aktivite durumundan faydalanarak bu iki evrenin birbirlerinden ayırt edilebilmesi ve uyku başlangıcının erken yakalanabilmesi için kısmi en küçük kareler regresyon yönteminden yararlanarak yapılan sınıflandırmanın başarısını ortaya koymaktır. Çalışmada karşılaştırma yapmak amacıyla aynı verilere k-en yakın komşuluk ve bayes sınıflandırma yöntemleri uygulanmış olup sonuçlar karşılaştırıldığında kısmi en küçük kareler regresyon yöntemi kullanılarak yapılan sınıflandırmanın % 90'lar civarında kabul edilebilir bir başarı oranı sağladığı ve zaman ile işlem yükü açısından daha avantajlı olduğu görülmüştür.

Çalışmada yedi farklı sağlıklı kişiye ait EEG kayıtları kullanılmış olup kullanılan EEG işaretleri physionet veri bankasındaki sleep-edf veri tabanında bulunan uyku skorlama çalışmalarına ait kayıtlardır. Bu kayıtlar normalizasyon ve filtreleme basamaklarından oluşan ön işleme tabi tutulmuş ve öznelik çıkarma amacıyla kayıtlara otoregresif modelleme yöntemi uygulanmıştır. Yapılan bu çalışmanın klinik uygulamalarda ve uyku ikaz sistemlerinde kullanılabilir olacağı düşünülmektedir

Uygulamalardaki tüm sonuçlar MATLAB programı ile elde edilmiştir.

**Anahtar Kelimeler:** EEG, Uyku, Sınıflandırma, Kısmi en küçük kareler regresyon, Otoregresif model, Bayes, k- en yakın komşuluk



Master Thesis

SUMMARY

THE CLASSIFICATION OF SLEEP EEG WITH PARTIAL LEAST SQUARES  
REGRESSION METHOD

Kübra EROĞLU

Karadeniz Technical University

The Graduate School of Science Electrical-Electronics Engineering Graduate Program

Supervisor: Assoc. Prof. Dr. Temel KAYIKÇIOĞLU

2013, 91 Pages

The aim of this study is to put forward the success of the classification carried out by using partial least squares regression in order to get sleep initiation early and to differentiate these two stages from each other by using brain activity situation observed during sleep and wake cycle observed in EEG sleep record. In order to make a comparison, k-nearest neighbor and bayes classification methods were applied with the same data and when the results were compared, classification carried out by using regression method was found 90 % successful and was seen more advantageous in terms of time and processing load. In the study, EEG records that belong to seven different healthy individuals were used and the EEG signs are gathered from the recordings that belong to sleep scoring studies present in sleep EDF database in physioBank. These records were exposed to pre-treatment composed of normalization and filter and autoregressive modelling method was used to extract the feature. It is supposed that this study can be used in clinical applications and in sleep warning systems.

All of the results in the study were gathered through MATLAB program.

**Key Words:** EEG, Sleep, Classification, Partial Least Squares, Autoregressive Model, Bayes, k-Nearest Neighbor

## ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1. Sinir hücrelerinin oluşturduğu aktivitenin yüzeyden ölçülmesi .....	4
Şekil 2. n-kanallı bir EEG ölçüm sisteminin blok diyagramı .....	5
Şekil 3. Beyinin farklı bölgeleri (frontal lob, parietal lob, oksipital lob, temporal lob).....	6
Şekil 4. Uluslar arası 10-20 sistemine göre elektrot yerleşimi, A: kafanın sol yandan görünümü, B: kafanın üstten görünümü, C: tüm elektrot noktalarının genel görünümü .....	7
Şekil 5. A: bipolar, B: unipolar ölçüm .....	8
Şekil 6. Göz açıp kapamanın alfa dalgası üzerindeki etkisi.....	9
Şekil 7. Farklı frekans bandındaki EEG dalgalarına ait örnekler .....	10
Şekil 8. Polisomnografi ölçümlerinde elektrot yerleşimi .....	12
Şekil 9. Farklı yaş gruplarında görülen farklı uyku dönemleri ve uyanıklığın zaman ile değişimi .....	14
Şekil 10. Farklı uyku basamaklarının dalga görünümleri .....	17
Şekil 11. Sağlıklı bir kişi için hipnogram örneği .....	18
Şekil 12. a) ham EEG verisi, b) normalize edilmiş EEG verisi .....	24
Şekil 13. Band geçiren filtrenin genlik-frekans tepkisi .....	25
Şekil 14. Butterworth filtre cevabı.....	25
Şekil 15. KEKK’de gizli değişken yapıları üzerine projeksiyonun temsili olarak gösterilmesi [50]. .....	47
Şekil 16. İki sınıfa ait olasılık yoğunluk fonksiyon grafiği .....	57
Şekil 17. Sağlıklı bir bireye ait beş dakikalık ham EEG kayıt örneği .....	59
Şekil 18. Veri bankasındaki bir kişiye ait hipnogram grafiği .....	60
Şekil 19. Sağlıklı bir bireyin uyanıklık ve uyku evrelerine ait epok örnekleri .....	61
Şekil 20. W (uyanıklık) evresine ait ortalama güç-frekans grafiği.....	62
Şekil 21. NREM1 evresine ait ortalama güç-frekans grafiği.....	63
Şekil 22. NREM2 evresine ait ortalama güç-frekans grafiği.....	63
Şekil 23. NREM3 evresine ait ortalama güç-frekans grafiği.....	63
Şekil 24. NREM4 evresine ait ortalama güç-frekans grafiği.....	64
Şekil 25. REM evresine ait ortalama güç-frekans grafiği.....	64
Şekil 26. Teta band aralığı için farklı filtre derecelerine ait (n5, n6, n7, n8, n9, n10) kazanç yanıtı ve butterworth genlik karakteristikleri .....	65

Şekil 27.	Alfa band aralığı için farklı filtre derecelerine ait (n5, n6, n7, n8, n9, n10) kazanç yanıtı ve butterworth genlik karakteristikleri .....	66
Şekil 28.	Beta band aralığı için farklı filtre derecelerine ait (n5, n6, n7, n8, n9, n10) kazanç yanıtı ve butterworth genlik karakteristikleri .....	66
Şekil 29.	EEG sinyali için AIC .....	67
Şekil 30.	Normalize edilmiş uyanıklık evresi EEG işaretine ait 30 saniyelik epok örneği .....	68
Şekil 31.	Normalize edilmiş NREM3 evresi EEG işaretine ait 30 saniyelik epok örneği .....	69
Şekil 32.	Belirlenen kurala göre verilerin etiketlenilmesi .....	72
Şekil 33.	Performans çalışması için şematik gösterim .....	80
Şekil 34.	Gerçek ve tahmin edilen veriler için uyanıklık/uyku grafikleri .....	81
Şekil 35.	Gerçek ve tahmin edilen veriler için uyanıklık/uyku grafikleri .....	81
Şekil 36.	Uyanıklık-uyku geçiş grafikleri .....	82
Şekil 37.	Gerçek ve tahmin edilen veriler için uyanıklık/uyku grafikleri .....	82
Şekil 38.	Uyanıklık-uyku geçiş grafikleri .....	83
Şekil 39.	Uyanıklık-uyku geçiş grafikleri .....	83
Şekil 40.	Uyanıklık-uyku geçiş grafikleri .....	84

## TABLolar DİZİNİ

	<b><u>Sayfa No</u></b>
Tablo 1. EEG işaretlerinin kapsadıkları bantlar.....	9
Tablo 2. Sağlıklı yetişkin bir bireye ait uyku evrelerinin yüzdesel dağılımı [10]. .....	17
Tablo 3. Veri tabanındaki her bireyin tüm uyku ve uyanıklık evrelerine ait epok sayıları.....	70
Tablo 4. Bir bireye ait eğitim amaçlı oluşturulan öz niteliklerin veri boyutları .....	71
Tablo 5. Bir bireye ait test amaçlı oluşturulan öz niteliklerin veri boyutları .....	71
Tablo 6. Bir bireye ait her band için tahminlenen sınıfların boyutları.....	73
Tablo 7. Belirlenen kritere göre olası durumlar ve tahmin edilen etiketler .....	73
Tablo 8. KEKKR yöntemi sınıflandırma başarı sonuçları.....	74
Tablo 9. Farklı epok uzunlukları kullanılarak elde edilen KEKKR yöntemi sınıflandırma başarı sonuçları.....	75
Tablo 10. Farklı epok uzunlukları kullanılarak elde edilen k-NN yöntemi sınıflandırma başarı sonuçları.....	76
Tablo 11. Farklı epok uzunlukları kullanılarak elde edilen bayes yöntemi sınıflandırma başarı sonuçları.....	78
Tablo 12. Her üç sınıflandırma yönteminde eğitim ve test aşamasında harcanan süreler .....	79
Tablo 13. Sürekli EEG kayıtları için KEKKR yöntemi ile sınıflandırma başarısı .....	84

## SEMBOLLER DİZİNİ

<b>A/D</b>	: Analog/sayısal dönüştürücü (Analog/digital)
$a_j^i$	: i. dereceden otoregresif katsayılar
<b>AIC</b>	: Akaike Bilgi Kriteri (Akaike Information criterion)
<b>AR</b>	: Otoregresif (Autoregressive)
<b>AASM</b>	: Amerikan Uyku Tıbbı Akademisi (American academy of Sleep Medicine)
<b>BT</b>	: Bilgisayarlı Tomografi
<b>CAP</b>	: Döngüsel Alternatif Desen (Cyclic Alternating Pattern)
<b>Cov(x, y)</b>	: x ile y arasındaki kovaryans
<b>dB</b>	: Desibel (Decibel)
<b>ECG</b>	: Elektrokardiyografi
<b>ECOG</b>	: Elektrokortiyografi
<b>EEG</b>	: Elektroensefelogram
<b>EKK</b>	: En Küçük Kareler
<b>EMG</b>	: Elektromiyografi
<b>EOG</b>	: Elektrookülografi
<b>E(x)</b>	: Beklenen değer
<b>E(x<sup>2</sup>)</b>	: Karesel beklendik değer
<b>FFT</b>	: Hızlı Fourier Dönüşümü (Fast Fourier Transform)
$f_s$	: Örnekleme frekansı
$F_p$	: Frontal kutup (frontal pole)
$g_i(x)$	: Diskriminant fonksiyonu
<b>Hz</b>	: Hertz
<b>MRG</b>	: Magnetik Rezonans Görüntüleme
<b>NREM</b>	: Hızlı Olmayan Göz Hareketleri (Non Rapid Eye Movement)
<b>KEKKR</b>	: Kısmi En Küçük Kareler Regresyon
$k_i$	: Yansıma katsayıları
<b>k-NN</b>	: k-En Yakın Komşuluk
<b>KEKK</b>	: Kısmi En küçük Kareler
<b>LVQ</b>	: Yapay sinir ağlarında vektör parçalama yöntemi (Learning Vector Quantization)

<b>MA</b>	: Yürüyen Ortalama (Moving Average)
<b>MUSIC</b>	: Çoklu Sinyal Sınıflandırma
<b>NIPALS</b>	: Doğrusal Olmayan Yinelemeli En Küçük Kareler (Non Linear Iterative Partial Least Squares)
<b>OLS</b>	: Klasik En Küçük Kareler
<b>ort</b>	: Ortalama
<b>REM</b>	: Hızlı Göz Hareketleri (Rapid Eye movement)
<b>RMSE</b>	: Hata Kareler Ortalamasının Karekökü (Root Mean Square Error)
<b>P(A)</b>	: A olayının önsel olasılığı
<b>P(B)</b>	: B olayının önsel olasılığı
<b>P(A/B)</b>	: B olayı bilindiğinde, A olayının olasılığı
<b>P(B/A)</b>	: A olayı bilindiğinde, B olayının olasılığı
<b>PzOz</b>	: Oksipital noktası referans olmak üzere parietal elektrot bölgesi
<b>SIMPLS</b>	: Kısmi en küçük kareler yönteminin istatistiksel olarak esinlenilmiş değişikliğinin basit bir uygulaması
<b>std</b>	: Standart sapma
<b>W</b>	: Uyanıklık (Wake)
$w_c$	: Filtre kesim frekansı
$X_n$	: X işaretinin normalizasyonu
$\bar{X}$	: X verisinin ortalaması
$\beta_{KEKK}$	: Kısmi en küçük kareler tahmin edicisi
$\beta_{OLS}$	: Klasik en küçük kareler tahmin edicisi
$\mu v$	: Mikro volt
$\alpha$	: Alfa dalgası
$\beta$	: Beta dalgası
$\delta$	: Delta dalgası
$\theta$	: Teta dalgası
$\gamma_p$	: Yansıma katsayısı
$\varepsilon$	: Hata vektörü
$\sigma^2$	: Varyans
$r_{xy}$	: Çapraz korelasyon
$\Sigma_i$	: Kovaryans matrisi
$\rho_{xy}$	: Korelasyon katsayısı

## 1. GENEL BİLGİLER

### 1.1. Giriş

Hayatın yaklaşık olarak üçte birlik bölümünü oluşturan uyku uzun yıllardan beri insanoğlunun yoğun olarak ilgilendiği konular arasında yer almış olup son zamanlarda bu konu hakkında elde edilen bilgiler önemli ölçüde artmıştır.

Uyku, sinir hücrelerinin meydana getirdiği elektriksel salınımların yapısal birleşimi ile karakterize edilebilen ve beyin aktivitesinin uyanıklık durumuna göre daha durağan olduğu bir durumdur [1]. Uyku süreci bilinenin aksine pasif bir dinlenme durumu olmayıp tüm vücudun yenilenmeye girdiği aktif bir süreçtir. Bugüne kadar yapılan birçok bilimsel araştırma gösteriyorki düzenli olarak gerçekleştirilen uyku tüm vücudun performansını artırırken başta merkezi sinir sistemi olmak üzere vücudun büyük bir kısmının yenilenmesini de sağlamaktadır. Uzun yıllar boyunca yapılan çalışmalara rağmen halen daha tamamen çözülemeyen uyku durumu, çalışmalarda elektroensefelogram (EEG)'nin kullanılmaya başlanması ile daha iyi anlaşılır hale gelmiştir.

Serebral kortekste bulunan ve beynin temel fonksiyonel birimi olan nöronlarda meydana gelen elektriksel aktivitelerin kaydedilmesi işlemine elektroensefalografi denilmektedir. Elde edilen kayıtlar ise EEG olarak adlandırılmaktadır. Bu yöntem ile beynin yapısal özelliklerinden çok fonksiyonel durumu hakkında bilgi alınmaktadır. Bu nedenle Magnetik Rezonans Görüntüleme (MRG), Bilgisayarlı Tomografi (BT) gibi yapısal incelemenin gerçekleştirildiği görüntüleme yöntemlerinde patolojik bir bulguya rastlanılmadığı durumlarda EEG'nin önemi daha da artmaktadır. EEG acısız, kolay bir inceleme yöntemi olduğu için başta epilepsi olmak üzere beyinle alakalı paraksomal hastalıklar, komalar, metabolik hastalıklar, kafa travmaları, tümör gibi birçok beyin rahatsızlığının teşhisinde, uykusuzluk, narkolepsi, insomnia, uyku apne sendromu gibi çeşitli uyku hastalıklarının tanınmasında ve uyku laboratuvarlarında gerçekleşen polisomnografi kayıtları ile uyku evrelerinin skorlanması ve skorlanan bu uyku evreleri ile uyku hastalıkları arasındaki ilişkinin ortaya konmasında, önemli bir yöntemdir.

Uyku çalışmalarında EEG'nin kullanılmaya başlaması ile çalışmalar hız kazanmış ve uyku süreci ile ilgili pek çok bilgi elde edilmiştir. Uykunun EEG ile analiz edilmesi

nörolog ve uyku uzmanlarına uykuya alakalı problem tanılarının konulmasında ve tanı sonrası tedavinin değerlendirilmesinde büyük bir kolaylık sağlamaktadır.

Bugüne kadar yapılan araştırmalar beyinsel aktivitenin uyanıklık ve uyku halinde farklılıklar gösterdiğini ortaya koymuştur. EEG işaretleri beynin fiziksel ve zihinsel aktivite durumuna göre dört farklı frekans bandına ait bileşenler ( $\delta$ ,  $\theta$ ,  $\alpha$ ,  $\beta$ ) içermektedir. Uyku EEG'sinde bu bileşenler incelenerek beynin durumu hakkında yorum yapılabilmektedir.

Bu çalışmada amaçlanan EEG kayıtlarında uyku ile uyanıklık evrelerinde gözlemlenen beyinsel aktivite durumundan faydalanarak uyku, uyanıklık ve bu iki durum arası geçiş sürecinde meydana gelen fizyolojik değişimlerin incelenerek uyku ile uyanıklık durumunun ayırt edilebilmesi ve uyku başlangıcının erken yakalanabilmesidir. Çalışmada kullanılan EEG kayıtları Physionet veri bankasından elde edilmiştir. Sağlık problemi olmayan yedi farklı kişiye ait EEG işaretleri öncelikle belirli ön işlemlere tabi tutulmuş ardından parametrik bir yöntem olan Otoregresif (Autoregressive-AR) model kullanılarak bu kayıtlardan uyku ve uyanıklık dönemlerine ait öznitelik vektörleri elde edilmiştir. Elde edilen öznitelik vektörlerinin uyku ve uyanıklık olmak üzere sınıf tahminlerinin yapılabilmesi için Bayes, k-En Yakın komşuluk (k-NN) ve Kısmi En Küçük Kareler Regresyon (Partial Least Squares Regression-KEKKR) yöntemleri kullanılmıştır. Tahminlenen sınıfların gerçek sınıflar ile karşılaştırılması sonucunda KEKKR yönteminin uyku ile uyanık durumunun ayırt edilmesinde yüksek başarı oranı sağladığı gözlemlenmiş olup ayrıca zaman ve işlem yükü açısından bu yöntemin daha avantajlı olduğu ortaya konulmuştur.

Tipik bir uyku süreci uyku ve uyanıklık döngüleri arasında gerçekleşen geçişlerden oluşmaktadır. Uzmanlar EEG işaretlerini inceleyerek kişinin uyku ve uyanıklık mekanizmalarını açıklayabilmekte fakat göz ile yapılan bu teşhisler hatalar içermekte olup zamansal anlamda uzun sürmektedir. Sonuç olarak uzmanların göz ile yaptıkları teşhisler bu yöntem ile daha az hata içererek daha kolay şekilde gerçekleştirilebilir.

Elde edilen sonuçların uyku analiz çalışmalarına ışık tutacağı düşünülmektedir. Ayrıca kullanılan yöntemin uyku başlangıcı sayılan yorgunluk durumunun erken belirlenmesine katkı sağlayarak sürücü güvenliğinin sağlanması amacıyla güçlü ve güvenilir yorgunluk ikaz araçları geliştirme çalışmalarında da kullanılabileceği düşünülmektedir.



Bu tez çalışmasında ilk olarak uyku EEG'sinin öneminden bahsedilip EEG ve uyku hakkında genel bilgiler verilmiştir. Sonraki bölümlerde ise uyku analizi için kullanılacak yöntemler ayrıntılı biçimde incelenip ardından bahsedilen yöntemlerin veri kümelerine uygulanması sonucunda elde edilen deneysel sonuçlar verilerek bu sonuçlar hakkında tartışılmıştır.

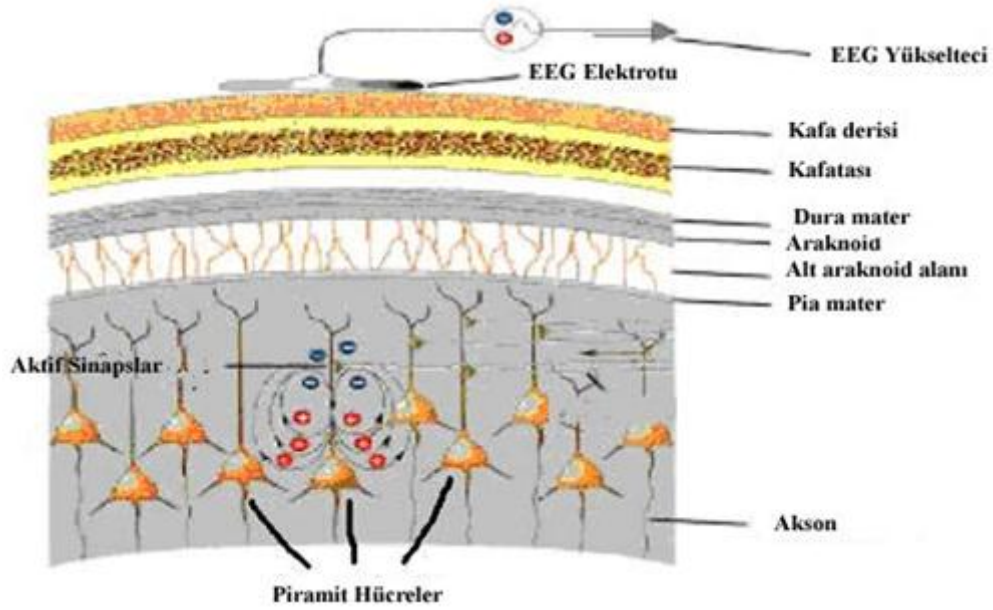
## 1.2. EEG

İnsan vücudunun en karmaşık yapısına sahip organlarından biri olan beyin sinirsel aktiviteler sonucunda kendiliğinden oluşan, düşük genlikli biyoelektriksel işaretler üretmektedir. Kafatası çevresine yerleştirilmiş elektrotlar aracılığı ile algılanan bu işaretler EEG işaretleri olarak adlandırılırlar. EEG işaretleri ilk olarak 1929 yılında Hans Berger tarafından kafatası üzerine yerleştirilen elektrotlar ve bu elektrotlara bağlı bir galvanometre aracılığı ile ortaya konulmuştur [2]. 1930 yılında ise çalışmasını geliştiren Berger EEG işaretlerinin gözün açılıp kapanma hareketi ile değiştiğini ortaya koymuştur. Adrian ve Matthews ise 1934'de EEG işaretlerinin elektrotlar aracılığı ile alınıp kuvvetlendirilerek kaydedilmesini sağlamışlardır. Beynin fonksiyonları ile alakalı yüksek oranda bilgi barındıran EEG işaretleri tıbbi teşhis ve araştırmalarda hastaya acı vermeyen ve maliyeti düşük bir yöntem olduğu için nöroloji klinikleri, anestezi derinliğinin tespiti, beyin bilgisayar arayüz çalışmaları, uyku analizi, ikaz sistemleri (otomotiv), epilepsi tedavisi gibi birçok alanda uzmanlar tarafından sıklıkla kullanılmaktadır.

EEG işaretleri insan zihinsel aktivite durumuna göre örneğin düşünme, oksijen eksikliği, uyku, uyanıklık gibi beynin farklı faaliyetleri sırasında değişim göstermektedirler. EEG işaretlerinin ana frekansları ile bu gibi zihinsel faaliyetler yakın ilişki içerisinde dirler. Düşük genlikte olan EEG işaretleri, güçlü arkaplan gürültüsüne sahip, durağan olmayan rastgele işaretler olup, beynin kompleks yapısı ve yönetmiş olduğu karmaşık işlemlerden ötürü uzmanlar tarafından görsel olarak yorumlanması güç işaretlerdir. Bu sebepten ötürü bu işaretlerin objektif şekilde analizlerini gerçekleştirebilmek için çeşitli yöntem ve bilgisayar teknikleri geliştirilmiştir. Geliştirilen tekniklerin bir kısmı EEG işaretlerini istatistiksel açıdan incelerken bir kısmı ise bu işaretlerde bulunan özel dalga şekillerinden faydalanmaktadırlar.

### 1.2.1. EEG İşaretlerinin Ölçülmesi

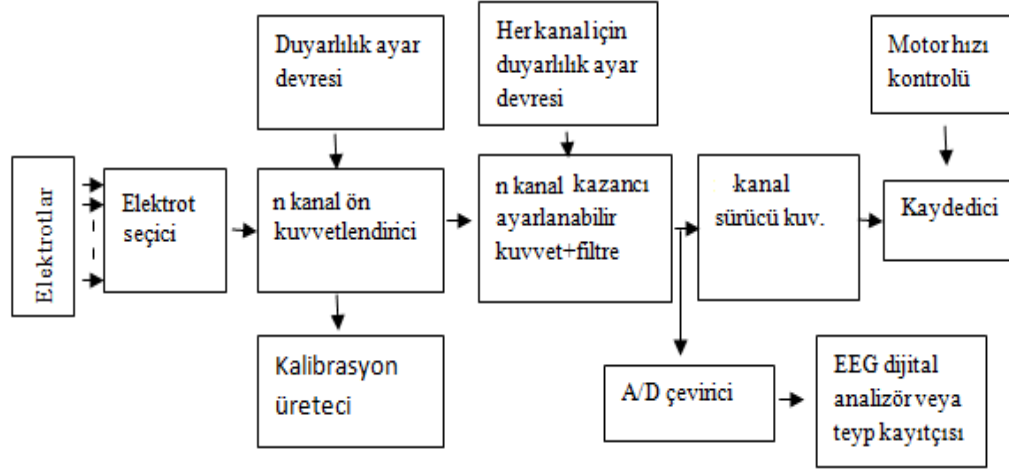
İnsan kafatası çok korunaklı bir yapıya sahip olduğundan beyinsel aktivite doğrudan beyin yüzeyinden gözlemlenememektedir. Bu sebeple kafa yüzeyine yerleştirilen elektrotlar aracılığı ile elde edilen veriler kuvvetlendirilerek EEG ölçümü yapılmaktadır. Kafa üzerinden yapılan bu ölçümler daha altta bulunan geniş bir bölgeden elde edilen gerilimlerin toplamı şeklindedir. Sinir hücrelerinin elektriksel aktivitesini gösteren EEG işaretlerinin elektrotlar aracılığı ile yüzeyden ölçülmesi Şekil 1 ile gösterilmektedir.



Şekil 1. Sinir hücrelerinin oluşturduğu aktivitenin yüzeyden ölçülmesi

Kafatasına yerleştirilen elektrotlar aracılığıyla elde edilen EEG işaretleri osiloskop tipi gözlem aracı ve bir kayıtçı tipi araç ile kağıta çizdirilebilmektedirler. Kafatasının iyi bir iletken olmaması, kemik dokusu, beyin sıvısı ve yağ dokularının beyinsel işaretleri zayıflatmasından ötürü kayıt için kullanılan aracın giriş empedansı ve kazancının yüksek olması gerekmektedir. Beynin farklı bölgelerindeki farklı aktiviteleri aynı anda ölçebilmek için EEG cihazları çoklu kanallı olarak kullanılmaktadırlar. Kanal sayısının artması çözünürlüğüde arttırmaktadır. Çok kanallı EEG ölçüm sisteminin blok diyagramı Şekil 2’de özetlenmiştir. EEG sinyal genliklerinin çok düşük (mikro volt’lar seviyesinde) olmasından ötürü yüksek kazançlı, gürültü seviyesi düşük, yüksek giriş empedansına sahip (1Mohm’dan yüksek), ortak mod eleme oranı yüksek (100 decibel (dB)’in üzerinde)

diferansiyel kuvvetlendiricilerin kullanılması gerekmektedir. Blok diyagramda da görüldüğü üzere A/D (Analog/Digital) dönüştürücü ile sayısallaştırılan EEG çıkışı bir kaydedici veya teyp kayıtcısı ile kaydedilmektedir.



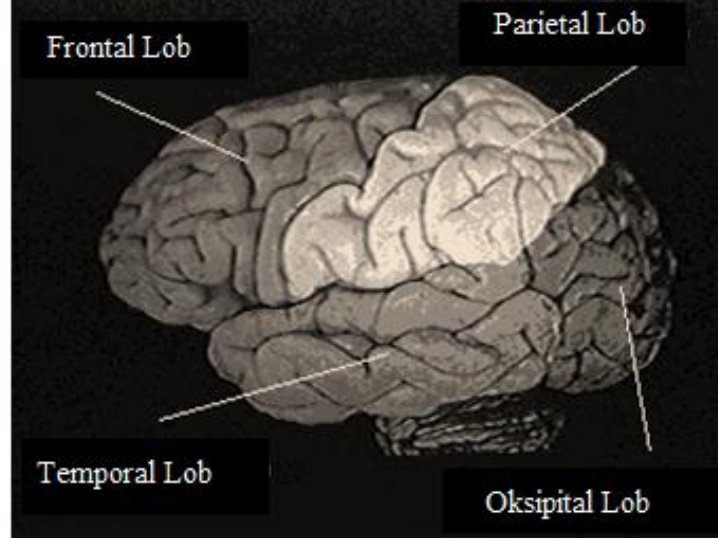
Şekil 2. n-kanallı bir EEG ölçüm sisteminin blok diyagramı [3].

EEG kaydı yapılırken iyi bir ölçüm yapılabilmesi için dışarıdan veya vücudun kendisinden ötürü oluşabilecek gürültüler ve olumsuz etkiler konusunda önlemler alınmalıdır. Ölçüm sonunda elde edilen EEG işaretlerinin oluşturmuş oldukları desenler işaret kaydının yapıldığı bölgeye, kayıt biçimine, beynin psikolojik ve fizyolojik durumlarına önemli ölçüde bağlı olup bu durumlara göre farklılıklar göstermektedir.

### 1.2.2. EEG Elektrotlarının Bağlantı Şekilleri

EEG kayıtları genel olarak kafatası üzerindeki belirli noktalara yerleştirilmiş olan elektrotlar aracılığıyla gerçekleştirilmektedir. Elektrotların kafatası üzerine yerleştirilmesinde kullanılan standart ise Uluslararası EEG Federasyonu Birliği (International Federation of Societies) tarafından belirlenmiş bulunan 10-20 elektrot sistemidir. Bu sistemde 19 aktif elektrot ve 1 adet kulak memesine yerleştirilen toprak elektrotu bulunmaktadır.

Beyindeki temel fonksiyon birimleri olan nöronlar serebral korteks üzerinde bulunmaktadır. Serebral korteks ise çeşitli beyin fonksiyonlarından sorumlu farklı bölgelere ayrılmaktadır. Bu bölgeler Şekil 3 ile gösterilmiştir.



Şekil 3. Beyinin farklı bölgeleri (frontal lob, parietal lob, oksipital lob, temporal lob)

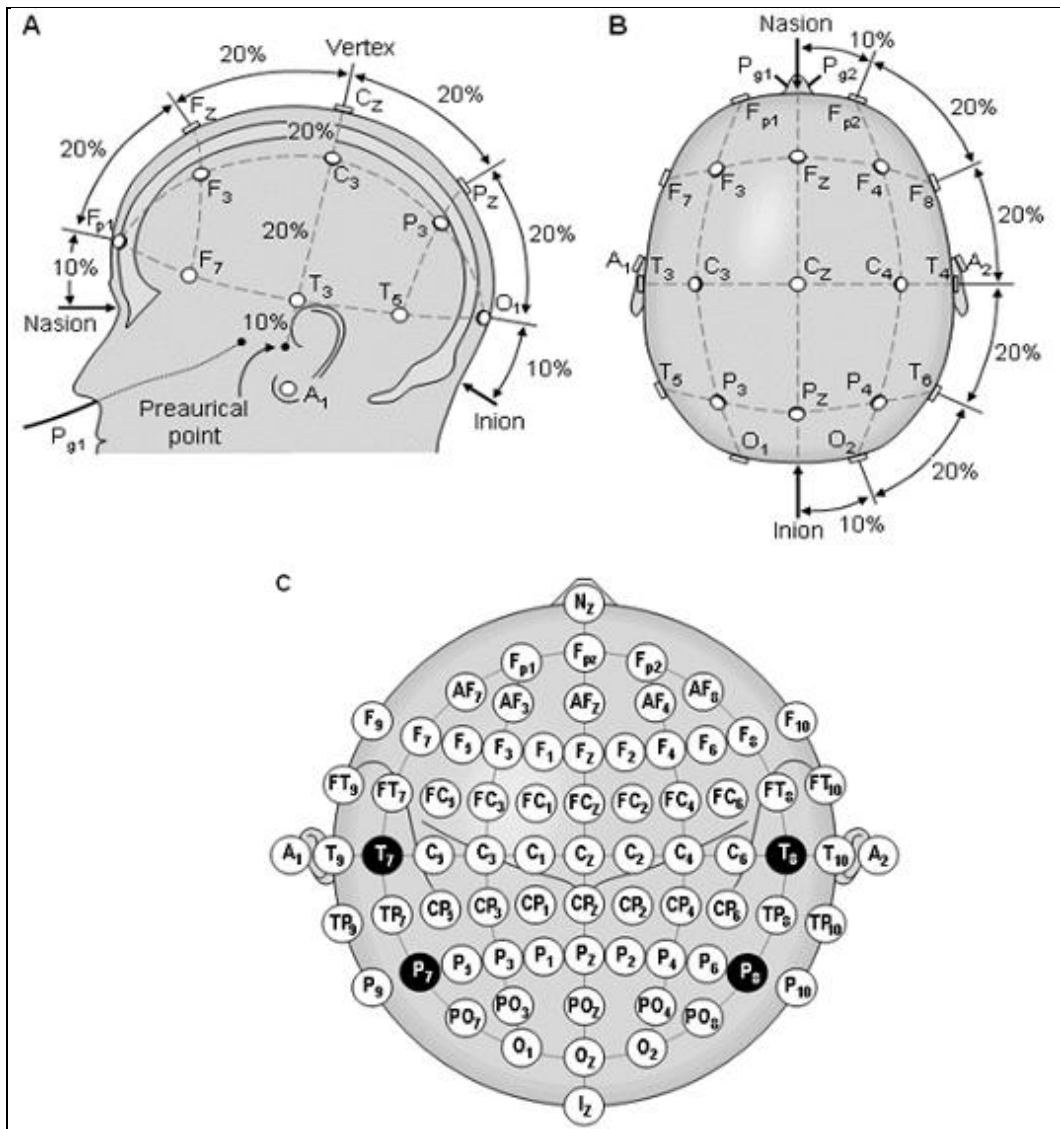
İşaret kaydı için kullanılacak olan elektrotların isimlendirilmesinde bu bölge isimlerinin baş harflerinde faydalanılmaktadır. Bu isimlendirme yapılırken bir harf ve bu harfin yanında bulunan küçük harf veya sayı kullanılır. Bu harf ve sayıların anlamı şu şekildedir;

- F: frontal (frontal)
- C: santral (central)
- O: oksipital (occipital)
- P: parietal (parietal)
- T: temporal (temporal)
- Fp: frontal kutup (frontal pole)
- A: sağ ve sol kulak (aurikuler)

Harflerin yanında yer alan daha küçük boyuttaki sayı ve harflerden tek sayılar sol yarım küreye, çift sayılar ise sağ yarım küreye yerleşecek olan elektrotları temsil ederken, küçük z harfi ise kafatasının orta hat bölgesini belirtir. Bunun yanında iki harfin birleşmesinden oluşan elektrot isimlendirmesi ise temel elektrot bölgeleri arasında kalan elektrot noktalarını ifade etmektedir.

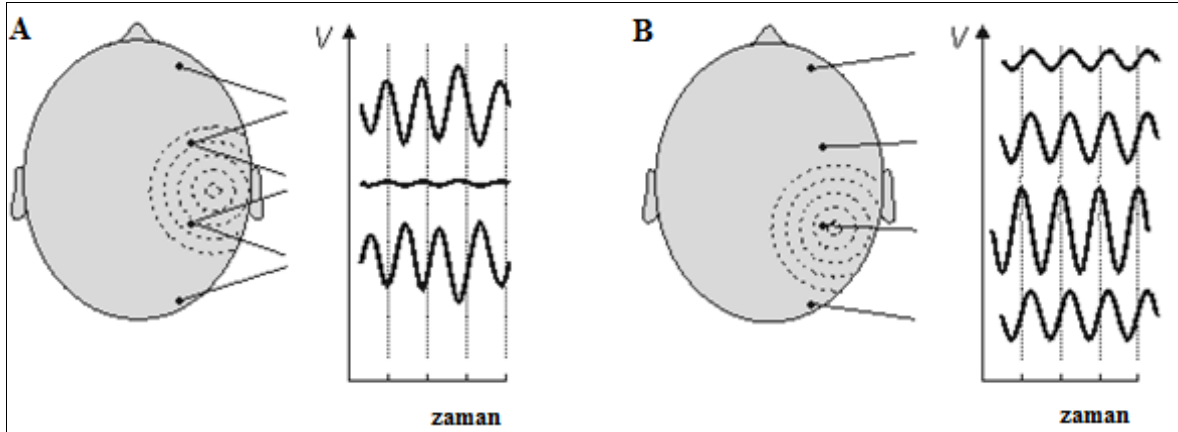
Elektrotların yerlerini belirleyebilmek için burnun alın ile birleştiği kısmı olan nasion, başın arka tarafındaki kısmında bulunan ve kafatasının arkadan dışarıya doğru çıkık kısmı olan inion, sağ kulak ve sol kulak kanalları önünde yer alan çukurluklar olmak üzere dört anatomik yer işareti kullanılır. Bu bölgeler kafatası üzerinde belirlendikten sonra

nasion ile inion noktaları arasındaki mesafe ölçülür ve iki nokta arasındaki hat hesaplanan toplam mesafenin sırasıyla %10, %20, %20, %20, %20, %10'i olacak şekilde aralıklara bölünerek bulunan noktalara elektrotlar yerleştirilir. Geriye kalan elektrotlar ise bulunan elektrot noktaları ile bir daire oluşturacak şekilde düzenlenir. Bu yerleştirme sistemi sayesinde farklı kişilerden farklı yerlerde kaydedilen EEG kayıtlarının karşılaştırılabilmesi sağlanabilir. Ayrıca daha fazla elektrot yerleştirilmesi yapılmak istenirse, bulunan elektrot noktaları arasındaki mesafeler tekrar bölünerek yeni elektrot noktaları oluşturulabilir. 10-20 sistemine göre elektrotların kafatası üzerine yerleştirilmesi Şekil 4 ile gösterilmiştir.



Şekil 4. Uluslararası 10-20 sistemine göre elektrot yerleşimi, A: kafanın sol yandan görünümü, B: kafanın üstten görünümü, C: tüm elektrot noktalarının genel görünümü [4].

Elektrot bağlantısı unipolar ve bipolar bağlantı olmak üzere iki şekilde yapılmaktadır. Unipolar bağlantıda merkezi sinir sisteminden uzak olan bir referans noktası seçilerek belirlenen elektrot bölgesinin bu referans noktasına göre potansiyel değişimi kaydedilir. Bipolar bağlantı şeklinde ise birbirine yakın iki referans bölgesi arasındaki potansiyel fark kaydedilir. Her iki bağlantı türünde Şekil 5 ile gösterilmiştir.



Şekil 5. A: bipolar, B: unipolar ölçüm [4].

Unipolar ölçüm kuvvetsizde olsa uzak bölgelerdeki aktiviteleri de içerdiğinden ötürü bipolar ölçüm unipolar ölçüme göre daha iyi sonuç vermektedir. Bu iki ölçüm şekline ek olarak bir de elektrokortiyografi (ECOG) adı verilen beyin kabuğuna elektrot dokundurulması ya da miyoelektrotlarla beynin derinliklerine inilerek gerçekleştirilen EEG sinyal kayıt şeklide mevcuttur. Bu kayıt biçimi ile EEG işaretlerini genliklerinin 3-4 katı büyüklükte olacak şekilde elde etmek mümkün olmaktadır.

### 1.2.3. EEG İşaretlerinin Kapsadığı Bantlar

EEG işaretleri geniş bir frekans spektrumuna sahip, tam olarak periyodik olmayan fakat ritmik sinyallerdir. EEG işaretlerinin sahip olduğu frekans aralığı beynin fonksiyonel aktivite durumuna göre 0,5-70 hertz (Hz), genlikleri ise 5-400  $\mu\text{V}$  (mikro volt) arasında değişmekte olup klinik ve fizyolojik ilgi ise 0,5-30 Hz arasına odaklanmıştır. Bu işaretlerin ana frekansları ile beyin fonksiyonları arasında sıkı bir ilişki bulunmaktadır. Beyin fonksiyonlarının aktivite düzeyleri arttıkça EEG dalgalarının frekansları da artmakta genlikleri ise azalmaktadır. EEG işaretleri içerdikleri dalgaların sahip olduğu baskın

frekanslara göre özel isimleri olan bantlara ayrılırlar. Bu bantların isimleri ile frekans ve genlik düzeyleri Tablo1 ile gösterilmiştir.

Tablo 1. EEG işaretlerinin kapsadıkları bantlar

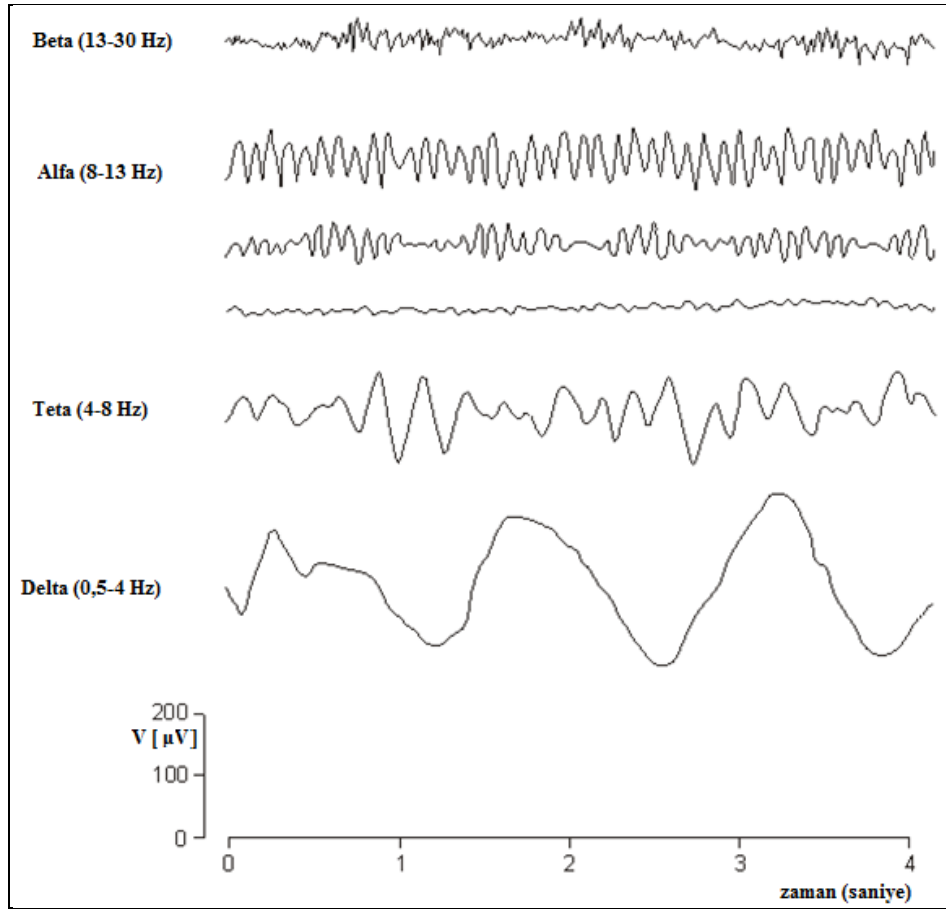
Dalga İsimleri	Frekans aralığı (Hz)	Genlik aralığı ( $\mu$ V)
Delta ( $\delta$ )	0,5-4	20-400
Teta ( $\theta$ )	4-8	5-100
Alfa ( $\alpha$ )	8-13	2-10
Beta ( $\beta$ )	13 ve üzeri	1-5

1. Delta Dalgaları ( $\delta$ ): Frekansları 0,5-4 Hz, genlikleri ise 20-400  $\mu$ V arasında değişmekte olan dalgalardır. Derin uyku, genel anestezi gibi beynin çok düşük aktivite gösterdiği durumlarda bu dalgalara rastlanılmaktadır.
2. Teta Dalgaları ( $\theta$ ): Frekansları 4-8 Hz, genlikleri ise 5-100  $\mu$ V arasında değişmekte olup orta derinlikte anestezi, rüyalı uyku, stres gibi beynin düşük aktivite gösterdiği durumlarda bu dalgalar ile karşılaşılır.
3. Alfa Dalgaları ( $\alpha$ ): Frekans aralıkları 8-13 Hz, genlik aralıkları ise 2-10  $\mu$ V olup uyanık bireylerin zihinsel ve fiziksel bakımdan tamamen dinlenme halinde olduğu, dış uyaranların bulunmadığı, gözlerin kapalı olduğu hallerde ortaya çıkarlar. Uyku durumunda ise yok olurlar. Oksipital bölgelerde belirgin olarak görülürler. Alfa ritminin gözlerin açık ve kapalı olduğu durumlardaki değişim şekil 6 ile ifade edilmiştir.



Şekil 6. Göz açıp kapamanın alfa dalgası üzerindeki etkisi

4. Beta Dalgaları ( $\beta$ ): Frekansları 13 Hz'den yüksek olup genlikleri ise 1-5  $\mu$ V arasında değişmektedir. Zihinsel işlev, Odaklanmış dikkat durumu, duyuşsal enformasyon işleme ve hızlı göz hareketlerinin bulunduğu uyku evrelerinde görülmektedirler. Beynin en yüksek aktivite düzeyine karşılık gelmektedirler. Beynin frontal ve parietal bölgelerinde belirgin olarak gözlemlenirler. Bahsi geçen tüm dalga şekillerine ait EEG işaret örnekleri Şekil 7 ile gösterilmiştir.



Şekil 7. Farklı frekans bandındaki EEG dalgalarına ait örnekler

### 1.3. Uyku

Uyku insanın etrafında gelişen olaylara tepkisiz kalarak uyanıklıktan ayırt edilebilen ve tersinir olan davranışsal bir durumdur. Tersinir olabilme durumu uykuyu tepkisizlik halleri olan anestezi ve koma gibi durumlardan ayırt eder [1]. Uyku sırasında insan beyni uyanıklık durumuna göre görece daha kararlı olan çeşitli psikofizyolojik durumlara geçer. Birçok sinir merkezi inaktif olur ve böylece beyin daha az karmaşık bir sistem haline

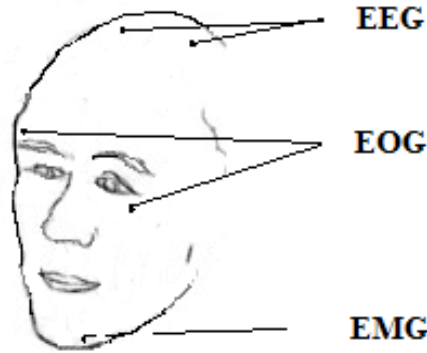


dönüşür. Bu durum aynı zamanda beyni matematiksel olarak modellenmeye daha uygun hale dönüştürür [5]. Tipik bir uyku süreci uyku ve uyanıklık döngüleri arasında gerçekleşen geçişlerden oluşmaktadır. Bu bakımdan doğal bir sirkadiyen ritm oluşturan uyku dinamik bir süreç meydana getirir. Temel işlevi zihinsel ve bedensel dinlenme olan uyku sırasında yaşamsal enerji depolanıp sinir sisteminin aktif olarak gelişimi sağlanır. Dokuların onarılması, kaynakların yenilenmesi ve yaşlanmanın gecikmesi, büyüme hormonlarının salgılanması gibi fizyolojik süreçlerin gerçekleşmesi, öğrenmemizi ve bellek oluşmasını sağlayan serebral değişikliklerin sağlanması yine uyku sırasında gerçekleşen önemli süreçlerdir. İnsan yaşamının yaklaşık üçte birlik kısmını oluşturan uyku kendi içerisinde Hızlı Göz Hareketleri (Rapid Eye Movement-REM) ve Hızlı Olmayan Göz Hareketleri (Non Rapid Eye movement-NREM) olmak üzere iki ayrı gruba ayrılmaktadır. NREM uykusu da kendi içerisinde NREM1, NREM2, NREM3 ve NREM4 olmak üzere dört evreye ayrılır. Uykunun farklı sınıfları olan REM, NREM ve NREM uykunun alt evreleri ile uyanıklık durumunun birbirlerine göre farklılık gösteren karakteristik özellikleri mevcuttur. Bu farklı özellikler uyku laboratuvarlarında elde edilen kayıtlarda uyku ile uyanıklık durumunun ayırt edilebilmesine olanak sağlamaktadır. Bahsi geçen bu özelliklere ilerleyen bölümlerde ayrıntısıyla değinilecektir.

### **1.3.1. Laboratuvarlardada Uyku Ölçümü**

Uyku süreciyle ilgili tanı, tedavi ve incelemelerin yapıldığı uyku laboratuvarlarında polisomnografik kayıtlar yapılarak uykuya ait temel veriler elde edilmektedir. Polisomnografi uyku sırasında uykunun yapısı ve fizyolojik değişimlerinin incelenmesidir. Bu sayede uykunun yapısı ve uykuda meydana gelen psikolojik ve patolojik değişimler uykunun evreleri ile ilişkilendirilebilmektedir [6]. Birçok fizyolojik parametrenin uzun süreli ve eşzamanlı olarak kaydını yapan cihazlara ise polisomnograf denilmektedir [7]. Polisomnografi kayıtları uyku laboratuvarlarında gece boyunca kişi uyurken elde edilmektedir. Kayıtların gerçekleştirildiği laboratuvar ortamları ısı, ses ve elektromanyetik izolasyonun sağlandığı ve mümkün olduğunca ev ortamına benzetilen yerlerdir. Uyku çalışmalarında kayıt alınacak kişiye çalışma yapılmadan önce uyku durumunu etkileyebilecek alkol, kafein ve türevi gibi maddeleri almaması ve günlük yaşam düzenini bozmaması konusunda bilgi verilir. Bu sayede kişinin daha sağlıklı bir uyku geçirmesi ve çalışma sonucunda da uykuyu etkileyebilecek maddelerden sakınıldığı için daha sağlıklı

kayıtlar elde edilmesi sağlanmış olur. Polisomnografi kayıtları ile beyinde gerçekleşen elektriksel aktiviteyi tespit eden EEG, kas hareketlerini tespit eden elektromiyografi (EMG), göz kası hareketlerini tespit eden elektrookülografi (EOG), kalp aktivitesini algılayan elektrokardiyografi (ECG) kayıtları gerçekleştirilip elde edilen veriler sayısal olarak kaydedilip görüntülenebilmektedir. Ayrıca bu kayıtların yanında solunum eforu, hava yolu, oksijen saturasyonu gibi değişkenlerin kayıtlarıda yapılmaktadır [8]. Bu kayıtlardan EEG, EMG ve EOG'nin baş üzerine yerleştirilen elektrotlar aracılığı ile nasıl gerçekleştirildiği şekil 8 ile temsili olarak ifade edilmiştir. Göz çevresine yerleştirilen elektrotlar uyku sırasında göz hareketlerinde meydana gelen voltaj değişikliğini ölçen EOG için, çene altına yerleştirilen elektrotlar uyku sırasında yüzde meydana gelen dramatik değişimleri ölçen EMG için, baş üzerine yerleştirilen elektrotlar ise beyinsel aktiviteyi ölçen EEG için kullanılmaktadırlar.



Şekil 8. Polisomnografi ölçümlerinde elektrot yerleşimi [5].

### 1.3.2. Uyku EEG'si

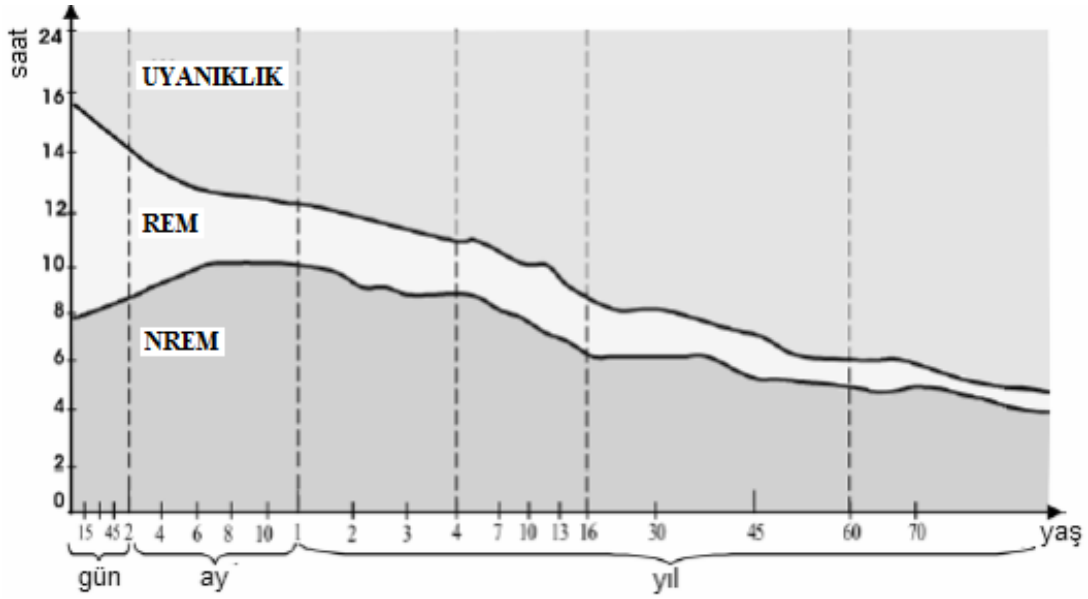
Uzun seneler boyunca yapılan çalışmalara rağmen halen daha tamamen çözülemeyen uyku durumu, çalışmalarda EEG'nin de kullanılmaya başlanması ile daha iyi anlaşılır hale gelmiştir. Uyku EEG kayıtları uyku apne sorunu, horlama, narkolepsi(aniden uyanma), insomnia, uyku bozuklukları gibi uyku hastalıklarında ve epilepsi gibi çeşitli hastalıkların tanısında ya da uyku analizlerinde kullanılmaktadır. Uyku EEG kayıtları gece veya gündüz kayıtları olarak 1-3 saat veya bütün gece süren kayıtlar şeklinde yapılmaktadır. Kısa süreli kayıtlarda kayıt sona erdikten sonra kişi uyandırılmakta, uzun süreli kayıtlarda ise kişinin gece boyunca sürekli olarak kayıt işlemleri gerçekleştirilmektedir. Kayıtlar sırasında

kişinin uyanık ve uykuda olma durumuna göre EEG kayıtlarındaki desenler farklılık göstermektedir. Bir kişinin EEG kayıtları incelenerek kişinin uyanık, uyku başlangıcında veya uykulu olup olmadığı eğer uyku halindeyse uykunun hangi evresinde bulunduğu saptanabilmektedir.

Uyku çalışmalarında uyku laboratuvarlarında gerçekleştirilen polisomnografi değerlendirmesi skorlama olarak isimlendirilen bir işlemle başlar. Skorlama işlemi ile EEG kayıtları üzerindeki uyku evrelerinin belirlenmesi sağlanmaktadır. Bu evrelerin belirlenmesinde EOG, ECG, EMG gibi farklı kayıt türlerinin verileride kullanılmakta olup tüm verilerin ortak olarak değerlendirilmesi sonucunda uyku evresi belirlenmektedir. Uyku evrelerinin skorlanması işleminde 30 saniyelik epoklar kullanılmaktadır. Her bir epok bir evre ile isimlendirilir. Eğer skorlama sırasında iki evre de aynı epokta yer alıyorsa, epokğun yarısından fazlası hangi evreye ait özellikler taşıyorsa epok o evre ile isimlendirilir.

### **1.3.3. Uyku Evreleri**

Daha öncede bahsedildiği gibi uyku yavaş dalga uykusu olan NREM ve hızlı dalga uykusu olan REM olmak üzere iki kısımdan oluşmaktadır. NREM uyku ise kendi içerisinde dört ayrı evreden oluşmaktadır. Uyanıklık, tamamen uyanıklık durumundan uyku haline kadar olan süreçtir. Uyku sürecindeki uyku döngüsü NREM uyku ile başlar ve bu süreçte yaklaşık olarak 90 dakikada bir NREM ve REM uykusu döngüsel olarak birbirini takip eder. Sağlıklı bir bireyde tüm gece uykusu boyunca bu döngüler 4-6 kez ardı sıra tekrarlanır. NREM uykusu uyku öncesindeki uyanıklık süresiyle ilişkili olup gecenin ilk üçte birine egemen iken REM uykusu ise gecenin son üçte birine hakimdir. Bu durum insan vücudunun biyolojik saati olan sirkadiyen ritim ile alakalıdır. NREM tüm uykunun %75-80'lik kısmını oluşturmaktadır. REM uyku ise %20-25'lik kısmını meydana getirir. Fakat bu durum şekil 9'da ifade edildiği gibi yaşa göre farklılıklar göstermektedir. Örneğin dünyaya yeni gelmiş bebeklerde uyanıklıktan uykuya geçiş sürecinde REM uykusu görülmekte olup birbirini takip eden NREM ve REM döngüsünün süreci ise 50- 60 dakikadır [9].



Şekil 9. Farklı yaş gruplarında görülen farklı uyku dönemleri ve uyanıklığın zaman ile değişimi [10].

NREM uyku yavaş dalga uykusu olarakta bilinmektedir. İlk NREM döngüsünün süresi yaklaşık olarak 70-100 dakika olup bu sürecin büyük bir kısmı NREM'in derin uyku kısmında (NREM3 ve NREM4) geçer. Bütün gece boyunca derin uyku dönemi azalırken REM uykusunun süresi ise giderek artar. Ayrıca bu durum yaş ile de değişmektedir. Yaş ilerledikçe yavaş dalga uyku sürecide kısaltmaya başlar [9]. NREM uykuda nabız ve kan basıncı düşer, kas tonusu azalır, beyin metabolizması en alt düzeydedir. Bu evrede hızlı göz hareketleri yoktur. NREM uykudan yaklaşık olarak 90 dakika sonra ise REM uykusuna geçilir.

REM uyku yaklaşık 10-20 dakikalık bir periyottur. Gece boyunca toplam 4-6 döngü şeklinde tekrarlanır. Bu evrede vücut solunumu düzensizleşir ve kalp atışları hızlanır. Hızlı göz hareketleri uykunun REM döneminde gözlemlenir. Rüyalarda bu evrede görülür. Beyin dalgaları tıpkı uyanık bir insanın beyin dalgalarına benzer fakat vücut kasları ise gevşektir. Bu yüzden bu uykuya paradoksal uyku da denilmektedir. REM uyku evresindeki kişiyi uyandırmak çok zordur. Uyandırılacak olsa bile kişi bir süre paralizi hissedebilir. Bunun yanında kişinin kendiliğinden veya aniden uyanması ise kolaydır. NREM uykunun alt evreleri EEG'de meydana gelen değişimlere göre ayırt edilebilmektedir. NREM1'den NREM4'e doğru ilerlenildikçe uyku dahada derinleşmektedir. Sağlıklı bir bireydeki uyanıklık, NREM ve REM uykunun EEG, EOG, EMG kayıtlarında görülen özellikler şu şekildedir;

1. Uyanıklık (Wake-W) : EEG'de oksipital bölgeden elde edilen kayıtlarda 30 saniyelik epğun yarısından fazlası alfa ritmine ait ise o epok uyanıklık evresi olarak skorlanır. Gözlemlenebilir bir alfa ritminin olmadığı durumlarda ise aşağıda açıklanan durumlardan birinin varlığı o epğun uyanıklık olarak skorlanması için yeterli olacaktır;
  - EOG'de 0,5-2 Hz frekansında göz kırpma hareketlerinin varlığı.
  - EOG'de okuma göz hareketleri olarak adlandırılan ve yavaş fazlı göz hareketlerini takiben zıt yönde hızlı göz hareketlerinin gözlemlendiği okuma sırasında ortaya çıkan göz hareketi.
  - EMG'de çene kas tonusunun normal veya yüksek olmasıyla birlikte EOG'de konjuge hızlı göz hareketi.
2. NREM1: Uyanıklık ile uyku arasındaki geçiş dönemi olan bu evre uykunun başlangıç evresidir ve yaklaşık olarak 1-15 dakika sürer. Yüzeysel uykudur. Kişi gözleri kapalı ve gevşemiş bir haldedir. Bu evrede kişi uykudan rahatlıkla uyanabilir. EEG kayıtlarında bir epğun yarısından fazlası düşük genlikli, teta frekansında (4-7 Hz) karışık frekanslı aktiviteden oluşuyorsa ve bu aktivite alfa dalgalarının yerini alıyorsa o epok NREM1 olarak skorlanır. Eğer epokta görülebilir alfa dalgası yoksa aşağıdaki durumlardan birinin varlığı NREM1 skorlanması için yeterli olacaktır;
  - Uyanıklık evresiyle kıyaslandığında EEG bazal aktivitesinde 1Hz yavaşlama
  - Verteks dalgalarının varlığı. Verteks keskin dalgaları santral bölgelerde en fazla görülen, zemin ritminden ayırt edilebilen, 0,5 saniyeden daha kısa süreli keskin kontürlü dalgalarıdır.
  - Yavaş göz küresi hareketleri
3. NREM2: 20 dakika kadar süren bu evre gerçek uykunun ilk aşamasıdır. EEG'de bir epğun birinci yarısında ya da önceki epğun son yarısında aşağıdaki durumlardan her ikisi ya da en az birinin olması halinde epok NREM2 olarak skorlanır.
  - Uyku içiği varlığı
  - Arousal ile ilişkisi olmayan K-kompleks varlığı (K kompleks ile birlikte 1 saniyeden uzun arousal varlığında ise evre NREM1 olarak skorlanır)

Uyku içcikleri frekansları 11-16 Hz arasında değişen ve en az 0,5 saniye süren keskin dalgalarıdır. K-kompleks ise EEG zemin ritmi aktivitesinden rahatlıkla ayırt

edilebilen negatif dalganın ardından hemen gözüken pozitif keskin dalganın oluşturduğu keskin dalgalardır. Toplam süreleri en az 0,5 saniyedir. Arousal ise fizyolojik ve davranışsal bakımdan koruyucu bir durum olan uykudan uyanmadır. Tüm bu bahsedilen durumlar dışında eğer bir epokta K-Kompleks veya uyku içiği gözlemlenmiyor ve epoktaki EEG aktivitesi düşük voltajlı aktiviteden oluşuyorsa önceki epok incelenir, bir önceki epok NREM2 ise NREM2 evresi devam etmektedir. Ayrıca NREM2 evresinde EOG'de gözün sağa ve sola kayması şeklinde yavaş göz hareketleri gözlemlenebilir. EMG'de ise kas aktivitesi düşüktür. Kişiyi uyandırmak biraz daha zordur.

4. NREM3: Orta dereceli derin uyku evresidir. EEG'de belirgin olarak frontal bölge kayıtlarında bir epoktaki 0,5-2 Hz frekanslı yavaş dalga aktivitesi epokun %20 veya daha fazlasını meydana getiriyorsa o epok NREM3 olarak skorlanır. EOG'de yavaş göz hareketleri görülmez. EEG'de uyku içikleri görülebilir. Çene EMG'si değişkenlik gösterir. Bu evrede kan basıncı ve vücut sıcaklığı düşer. Kişinin uyandırılması zordur.
5. NREM4: Derin uyku evresidir. EEG kayıtlarında epokun yarısından fazlasını delta aktivitesi oluşturur. K-kompleks ve uyku içikleri kaybolur. EMG'de kas tonusu çok düşük olup EOG'de ise göz hareketleri gözlemlenmez. Uykuda konuşma, gezme, ıslatma gibi durumlar bu evrede gözlemlenir.
6. REM: Rüyalı uyku evresidir. EOG'de hızlı göz hareketleri gözlemlenir. EMG'de kas tonusu en düşük seviyededir. Çok kısa süreli ani artışlar görülebilir. EEG düşük genlikli ve karışık frekanslıdır.

Rechtschaffen ve Kales'in 1968 yılında editörlüklerini yaptıkları 12 kişilik araştırmacı grubu insan uyku evrelerinin standart terminoloji, teknik ve skorlama el kitabını hazırlamışlardır. Yakın tarihe kadar uyku evrelerinin skorlanması işleminde bu kitapta yer alan esaslara göre uyku NREM1, NREM2, NREM3, NREM4 ve REM olmak üzere beş evreden oluşmaktaydı. Fakat Amerikan Uyku Tıbbı Akademisi'nin (American Academy of Sleep Medicine-AASM) Dr. Iber Conrad başkanlığında oluşturdukları çalışma grubunun çalışmaları sonucunda uyku skorlanması konusunda yeni kurallar belirlenmiş ve 2007 yılında bu kurallar yayınlanmıştır. Günümüzde bu kurallar skorlama çalışmalarında esas alınmaktadır. Bu yeni kurallara gereğince NREM3 ve NREM4 tek bir çatı altında NREM3 olarak adlandırılmıştır. Böylece uyku NREM1, NREM2, NREM3 VE REM olmak üzere dört evreden oluşmaktadır. Yukarıda açıklanan uyku evrelerinin özellikleri bu yeni kurallara göre yazılmış olup NREM3 ve NREM4 tek bir isimle NREM3 olarak

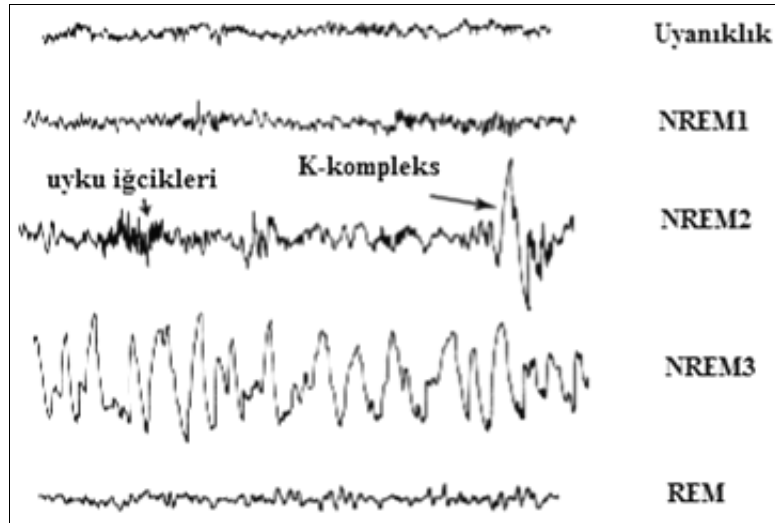
isimlendirilmektedir. Yapılan bu çalışmada kullanılan uyku EEG kayıtları Physionet veri tabanından alınmış olup eski kural sistemine göre isimlendirilme yapılmıştır. Fakat bu durum çalışmamızı etkilememektedir.

Sağlıklı bir yetişkin bireye ait uyku evrelerinin yüzdesel dağılımı tablo 2’de görülmektedir.

Tablo 2. Sağlıklı yetişkin bir bireye ait uyku evrelerinin yüzdesel dağılımı [10].

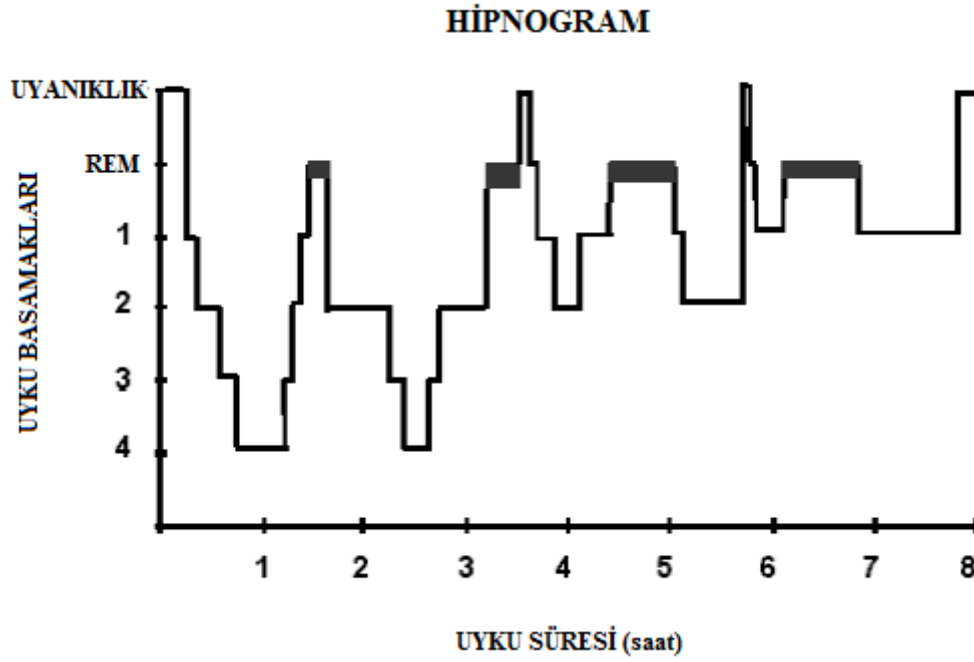
NREM1	% 5-10
NREM2	% 45-60
NREM3 ve NREM4	% 20-25
REM	%20-30

NREM, REM, uyanıklık dönemlerine ait dalga şekilleri ve NREM uykuda görülen uyku içcikleri ile K-kompleks yapıları ise şekil 10 ile gösterilmiştir.



Şekil 10. Farklı uyku basamaklarının dalga görünüşleri [5].

Uyku skortlama işleminde tüm uyku süresi boyunca skorlanmış olan uyku evrelerini gösteren histogram şeklindeki yapıya hipnogram denilmektedir. Şekil 11 ile bir hipnogram örneği gösterilmiştir. Bu yapı incelenerek skortlama işlemi sırasında süregelen zamanın hangi uyku evresine denk geldiği görülebilir.



Şekil 11. Sağlıklı bir kişi için hipnogram örneği

#### 1.4. Literatür Araştırması

Uyku konusunda yapılan çalışmalar EEG'nin bu çalışmalarda kullanılmasıyla beraber büyük bir hız kazanmıştır. EEG'nin uyku çalışmalarında kullanımı ile uyku ve uyanıklık sırasında beyinin elektriksel aktivitesinde meydana gelen farklı durumlar ortaya konulup incelenebilmiştir. Bu alanda yapılan ilk çalışmalar Hans Berger'e aittir. Daha sonrasında ise Loomis ve arkadaşları tarafından 1937 yılında uykunun evrelere ayrılması konusunda bir çalışma yapılmıştır. Uyku ve uyanıklık aşamalarının algılanabilmesi ve birbirlerinden ayırt edilebilmesi uyku çalışmalarında büyük bir önem taşımaktadır. Bu konuyla ilgili literatür araştırıldığında bir çok çalışma görülmektedir. Yakın dönem içerisinde yapılan çalışmalar incelenecek olursa;

Musa H. Asyalı ve diğerleri uyku derinliğini hesaplayabilmek için uyku durumundan uyanıklık durumuna ya da daha hafif uyku evresine geçiş anları olan ve 'Arousal' olarak adlandırılan geçiş evrelerini AR model tabanlı yöntem geliştirerek incelemişler. Çalışmada kafatasının iki ayrı bölgesinden (Oksipital, Santral) elde edilen EEG işaretleri kullanılmış olup bulunan bulgular sayesinde alfa ve beta bandındaki mutlak güç değerlerinin toplamının uyku derinliğini belirlemede bir belirteç olduğu gösterilmiştir [11].



Florian Chapotot ve Guillaume Becq çalışmalarında uyku evrelerinin, uyanıklık, NREM1, NREM2, NREM3, NREM4, REM ve hareket zamanı olmak üzere altı farklı evrede tam otomatik olarak sınıflandırılabilmesi için EMG, EOG ve dört kanallı EEG kayıtları kullanılmışlardır. Bu kayıtlardan elde edilen verilerden oluşturulan özellik vektörleri ise Yapay Sinir Ağları ve Basit Karar Kuralları birleştirilerek sınıflandırılması sağlanmıştır. Çalışma sonuçlarında uyanıklık, orta derinlikte uyku basamakları ve hareket zamanı %  $36 \pm 15$  sınıflandırma performansı ile buna karşın paradoksal uyku ve derin uyku ise % 82 sınıflandırma performansı ile sınıflandırılmıştır [12].

Anna Krakovska ve Kristina Mezeiova uyku basamaklarının otomatik olarak sınıflandırılmasına yönelik yaptıkları çalışmada 20 sağlıklı kişiden elde edilen polisomnografik kayıtlardan (EEG, EOG, ECG, EMG ) en iyi karakteristiğe sahip kümenin bulunması hedeflenmiş ve bu amaçla çeşitli kanal ve kanal kombinasyonlarına ait 74 ölçüm elde edilip hipnogram verileri ile karşılaştırılmıştır. Çalışma sonucunda 4-14 polisomnografik özelliğin bu amaç için yeterli olduğu ortaya konuldu [13].

Natheer Khasawneh ve diğerleri tarafından yapılan çalışmada ağ tabanlı bulanık çıkarım sistemi kullanılarak beş kanallı EEG, EOG ve EMG kayıtlarından elde edilen özellikler uyku basamaklarının sınıflandırılması için kullanılarak ortalama % 76.43 doğru tahminleme gerçekleştirildi [14].

Mourad Adnane ve diğerleri uyku-uyanıklık basamaklarını sınıflandırırken ECG kayıtlarından yararlandılar. Yaptıkları çalışmada Destek Vektör Makine yöntem tabanlı sınıflandırma sonucunda iki farklı sayıdaki özellik grubu için ortalama % 79.31 ve % 79.99 sınıflandırma doğruluğu elde ettiler [15].

Rajendra U. Acharya ve diğerleri altı farklı doğrusal olmayan parametre (Korelasyon boyutu, yaklaşık entropi, en büyük liapunov üstelleri, hurst bileşeni, fraktal boyutu, yineleme çizimi) kullanarak her bir uyku basamağı için beyin kortikal fonksiyonunu analiz ettiler. Uyanıklık basamağından uykunun dördüncü basamağına doğru doğrusal olmayan bu parametre değerlerinde bir azalma, uykunun beşinci basamağı olan REM evresinde ise yüksek korteks aktivitesinden dolayı artma gözlemlenildi [16].

Dilan Görür ve diğerleri uykunun NREM evresinde yoğun olarak gözlemlenen ve frekans özellikleri sayesinde diğer beyin aktivitelerinden ayrıştırılabilen uyku içcik aktivitelerinin yerlerini saptayabilmek için Kısa Zamanlı Fourier Dönüşüm yöntemi kullanarak elde ettikleri özneliklerin Çok Katmanlı Perseptron ve Destek Vektör Makine

yöntemleri ile sınıflandırılması sonucunda sırasıyla % 88.7 ve % 95.4 sınıflandırma performansı elde etmişlerdir [17].

Diego Alvarz Estevez ve Vicente Moret Bonillo gün içerisinde uykulu olma durumunun temel nedenlerinden olan uyku bölünmelerini yani arousalları polisomnografik kayıtlarda otomatik olarak belirlemek için bir metot geliştirdiler. Yaptıkları çalışmada kullandıkları algoritma ile EMG ve iki kanallı EEG kayıtlarından yararlanarak uyku bölünmelerinin belirlenebilmesi sağlandı. Yirmi hastadaki uyku bölünmeleri üzerinde yapılan çalışmada karşılaştırma sağlamak amacı ile doğrusal diskriminant, kuadratik diskriminant, destek vektör makine ve yapay sinir ağları yöntemleriyle sınıflandırma gerçekleştirildi. Yapay sinir ağları kullanılarak elde edilen duyarlılık ve özgüllük sonuçları (sırasıyla 0.86, 0.76 ) diğer yöntemlere göre daha iyi sonuçlar ortaya koydu [18].

Jo G. Han ve diğerleri EOG, EMG ve tek kanallı EEG kayıtlarından yararlanarak uyku durumunu uyanıklık, yüzeysel uyku, derin uyku ve REM evresi olmak üzere dört basamakta otomatik olarak sınıflandırabilen bir algoritma tasarladılar. Kullandıkları algoritma Fast Fourier Transform (FFT) ve Genetik Algoritma tabanlı olup uyku basamakları Bulanık Mantık yöntemi ile yaklaşık olarak % 84.6 doğrulukla sınıflandırılmıştır [19].

Salih Güneş ve diğerleri yapmış oldukları çalışmada EEG uyku sinyallerinin altı farklı evrede otomatik olarak sınıflandırılabilmesi için bu sinyaller 30 saniyelik epoklara ayrılmış ve her bir epok için FFT tabanlı Welch spektrum analizi uygulanarak EEG sinyallerine ait önemli özellikler elde edilmiştir. Bu özelliklerin sayısını azaltmak için bir takım istatistiksel yöntemler kullanılmış (minimum değer, maksimum değer, standart sapma, ortalama değer) ve elde edilen özellikler k-means algoritması ile ağırlıklandırılarak ön işleme alınmıştır. Ağırlıklandırılmış uyku basamaklarının k-NN sınıflandırma yöntemi ile sınıflandırılması sonucunda % 82.15 oranında başarı sağlanmıştır [20].

Sara Mariani ve diğerleri uyku EEG'inde gözlemlenen ve cyclic alternating pattern (CAP) olarak adlandırılan periyodik aktivitelere A fazı olarak karakterize edilen aktivite fazının otomatik olarak belirlenebilmesine yönelik bir çalışma gerçekleştirmişlerdir. A fazı EEG aktivitesinde gözlemlenen baskın delta dalgaları, hızlı göz hareketleri ve beta dalgalarının bulunma durumlarına göre belirlenmekte ve bu durumlara göre üç ayrı alt gruba ayrılmaktadır (A1, A2, A3). Çalışmada sekiz polisomnografik kayıt kullanılmış olup karşılaştırma yapmak amacıyla diskriminant sınıflayıcı, destek vektör makinesi, adaptif boosting ve denetimli yapay sinir ağları teknikleri kullanılarak elde edilen sınıflandırma

sonuçları görsel analiz sonuçlarıyla karşılaştırılmış ve en yüksek başarı doğrusal diskriminant tekniği sonucunda %84.9 doğruluk oranı ile elde edilmiştir [21].

Muhittin Bayram yapmış olduğu çalışmada EEG kayıtlarından elde ettiği verileri uyku, uyanıklık ve uyku ile uyanıklık arası geçiş dönemi kabul edilen uyuklama evresi olarak üç sınıfta sınıflayan bir algoritma tasarlamıştır. Bu çalışmada EEG kayıtları beş saniyelik epoklar halinde incelenmiş olup her bir epok dalgacık dönüşümü yöntemi ile dört alt frekans bandına ayrılarak (delta, teta, alfa, beta) elde edilen bölütlerden uyku, uyanıklık ve uyuklama karakteristiğine sahip bölütler seçilerek yapay sinir ağları yöntemi ile eğitilmiştir. Geliştirilen bu algoritma ile yüzlerce EEG bölütünün test edilmesi sonucunda düşük bir hata oranı ile EEG verileri sınıflandırılmıştır [22]. Aynı amaç ile Hatice Batar tarafından yapılan çalışmada ise yine Dalgacık Dönüşümü ve Yapay Sinir Ağları kullanılmış olup Learning Vector Quantization (LVQ) ağının kısa sürede öğrenebilme ve yüksek doğrulukta sınıflandırma yapabilme başarısı ortaya konulmuştur. Elde edilen başarı oranları uyanıklık, uyku ve uyuklama için sırasıyla % 99.4, %99.9 ve %99.8 şeklindedir [23].

Fazıl Duman ise uykunun ikinci evresinde gözlemlenen uyku içciklerinin kestirimini sağlayarak uyku skorlama çalışmalarında uzmanlara yardımcı olabilecek bir algoritma geliştirmiştir. Algoritmada içciklerin hem genlik hem de frekans analiz yöntemleriyle saptanması için dört farklı yöntem kullanılmıştır. Frekans analizi için Kısa Zamanlı Fourier Dönüşümü ve Çoklu Sinyal Sınıflandırma yöntemi (MUSIC), genlik analizi için Teager Enerji Operatörü, uyku içciklerinin bulunduğu frekans bandının tespiti için ise Dalgacık Dönüşümü kullanılmıştır. Çalışma sonucunda uyku içciklerinin kestirimi % 94 doğruluk oranıyla gerçekleştirilmiştir [1]. Uyku içciklerinin kestirimi ile ilgili Sinem Özkara Torun tarafından yapılan çalışmada ise içciklerin kestirimi için ardışık tahminleme yöntemlerinden yararlanılabileceği fikri ortaya konulmuştur [24].

## 2. YAPILAN ÇALIŞMALAR

### 2.1. Materyal

Çalışmada kullanılan kullanılan EEG işaretleri physionet veri bankasında bulunan sleep-edf veritabanından elde edilmiş olup bu veri tabanına '<http://www.physionet.org/>' internet adresinden ulaşılabilmektedir.

Yapılan çalışmada yukarıda bahsedilen veri tabanında bulunan 'sc\*.rec' ve 'st\*.rec' dosya kayıt isimlerine sahip gönüllü yedi sağlıklı kişiye ait ham EEG uyku kayıtları ve hipnogram verileri kullanılmıştır. Bu kişiler 21-35 yaş aralığında olup hiçbir medikal destek almamaktadırlar. Kayıtlar her bir kişi için kafatasının FpzCz ve PzOz olmak üzere iki bölgeden de ayrı olarak elde edilmiş olup 100 Hz örnekleme frekansına sahiptirler. Bu verilere ek olarak sitede EOG, EMG kayıtları da bulunmaktadır. Yapılan bu çalışmada uyanıklık ve uyanıklık-uyku arası geçiş evresini iyi temsil ettiği düşünülen elektrod pozisyonu olan PzOz (Kafatasının orta hat kısmındaki oksipital noktası referans olmak üzere yine orta hat üzerindeki parietal elektrot) noktasından alınan kayıtlar kullanılmıştır. Daha önceki bölümlerde de belirtildiği gibi uyanıklık durumundan uykunun ilk evresine geçiş aşamasında alfa dalgaları yerini teta dalgalarına bırakmaktaydı ve alfa dalgalarının en iyi gözlemlendiği bölge ise parietal bölgedir.

'sc\*.rec' kayıtları 1989 tarihinde sağlıklı üç bireyden günlük yaşamları sırasında 24 saat süreyle geliştirilmiş teyp kayıtları kullanılarak elde edilen kayıtlardır. 'st\*.rec' kayıtları ise 1994 tarihinde hafif uykuya dalma güçlüğü çekemeleri dışında sağlık problemi bulunmayan kişilerden gece boyunca hastanede minyatür telemetri sistemiyle ve iyi sinyal kalitesiyle elde edilen kayıtlardır. Veri tabanında bulunan hipnogram kayıtlarındaki uyku basamakları W, 1, 2, 3, 4, R, M ve skorlanmamış epoklardan oluşmaktadır. W, uyanıklık durumunu, R, REM evreyi, M, hareket halini, 1-4 arası numaralandırma ise sırasıyla uykunun dört evresini (NREM1, NREM2, NREM3, NREM4) belirtmektedir.

## 2.2. Yöntem

### 2.2.1. Ön İşleme

Yapılan çalışmada kullanılan tüm programlar MATLAB'da oluşturulmuş olup veri tabanında 'rec.' dosyası şeklinde kayıtlı olan EEG verilerinin 'mat.' formatına çevrilmesi için EEGlab (EEG sinyallerinin işlenmesi için interaktif araç kutusu)'dan faydalanılmıştır. Kullanılan EEG kayıtları 30 saniye uzunluğunda epoklara ayrılmış olup 100 Hz örnekleme frekansı ile örneklenmiş işaretlerdir. Bu çalışmada başarı karşılaştırması yapabilmek amacıyla 30 saniyelik epoklar dışında kullanılan algoritma 15, 10 ve 5 saniyelik epoklara ayrılmış EEG kayıtları için de tekrarlanılmıştır. Çalışmada EEG kayıtlarındaki her bir uyanıklık ve uyku evresi (W, NREM1, NREM2, NREM3, NREM4, REM) verilerini ayrı ayrı elde edebilmek için hipnogramdan yararlanılarak her kişi için kendi uyku dönemlerine ait beş ayrı evre ve uyanıklık verileri oluşturuldu.

#### 2.2.1.1. Normalizasyon

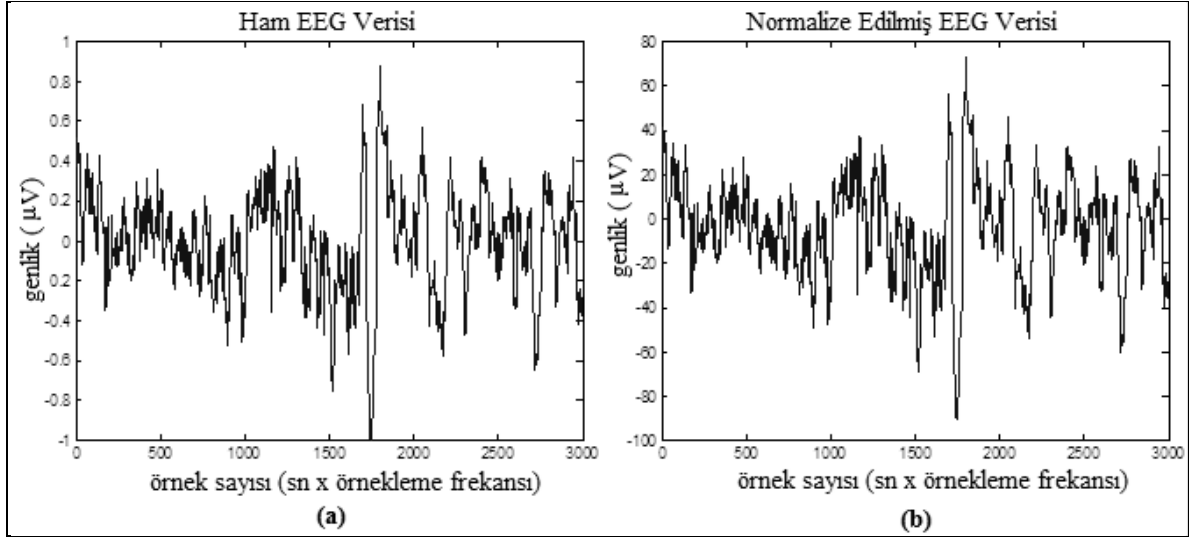
EEG işaretleri çok fazla gürültü barındıran işaretlerdir. Bu sebeple bu işaretlerde genlik farklılıkları ortaya çıkabilmektedir. Normalizasyon işleminin amacı veriler arasındaki farklılığın çok fazla olduğu hallerde verileri tek bir düzen içerisine sokabilmektir. Bir başka kullanım alanı ise ölçekleme sistemlerindeki verilerin karşılaştırılabilmesidir.

Normalizasyon işlemi ile ( $m \times p$ ) boyutlu veri kümesi bir uzaydan başka bir uzaya taşınır. Bu taşıma ile yeni maksimum ve minimum noktaları oluşur fakat ( $m \times p$ ) boyutu değişmez [25]. Normalizasyon işlemi için farklı yöntemler olmakla birlikte bu çalışmada eşitlik (1) ile ifade edilen normalizasyon işlemi ile EEG verileri ön işlemeye tabi tutulmaktadır. Bu işlem sonucunda EEG ham verileri  $[-1,1]$  aralığında ölçeklenmektedir.

$$X_n = \frac{x - \bar{x}}{\max|x - \bar{x}|} \quad (1)$$

Eşitlikteki  $X$  ham EEG verisini,  $X_n$  ham EEG işaretinin normalizasyonu sonucu elde edilen EEG verisini,  $\bar{x}$ ,  $X$  verisinin ortalamasını,  $\max$  ise en yüksek değeri ifade

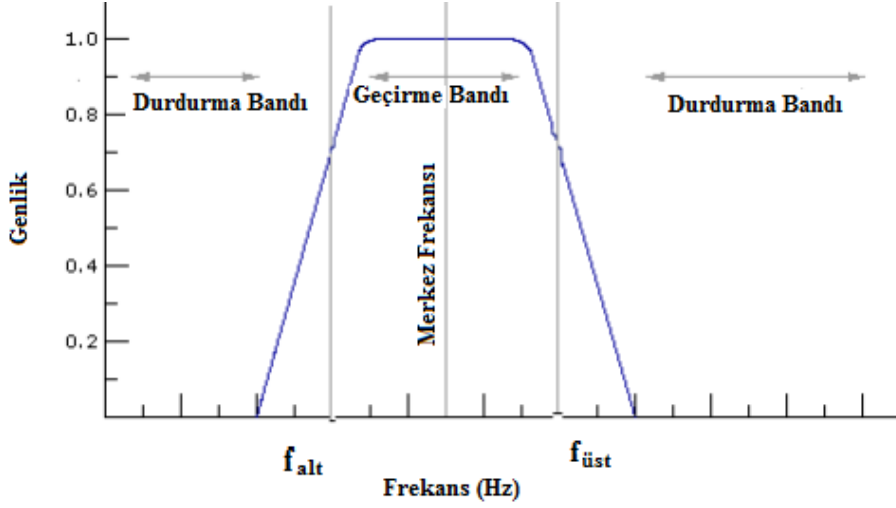
etmektedir. Şekil 12 ile verilen örnekte ham EEG verisine normalizasyon işlemi uygulandıktan sonra genlik değerlerinin  $[-1,1]$  aralığında değişimi görülmektedir.



Şekil 12. a) ham EEG verisi, b) normalize edilmiş EEG verisi

### 2.2.1.2. Filtreleme

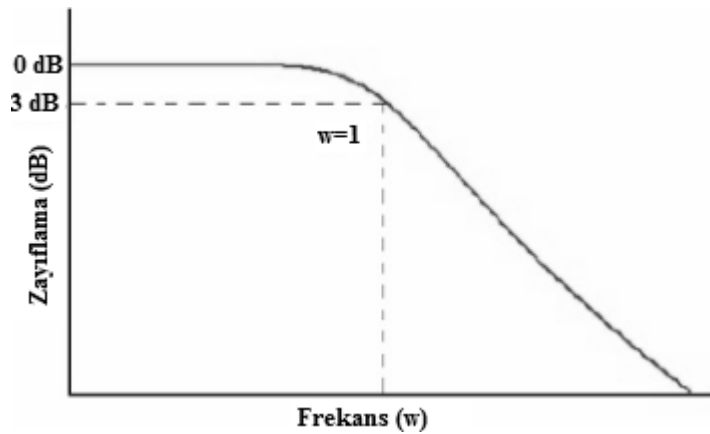
Genel tanımıyla filtreleme, işaretlerin bazı bileşenlerinin bastırılması veya istenilen kısımlarının aktarılması işlemleri olarak tanımlanabilir. Genel olarak orijinal işarete karışmış olan gürültüleri süzmek ya da istenmeyen bileşenlerin oluşturdukları etkileri en aza indirmek amacı ile kullanılırlar. Bu çalışmada ön işlemenin ikinci aşaması olarak normalize edilmiş ham EEG işaretlerinden çalışmanın amacına uygun olarak odaklanılması gereken frekans bandlarının elde edilebilmesi için incelenmek istenen her bir frekans aralığına ait band geçiren filtre tasarımı gerçekleştirilmiştir. Band geçiren filtreler belirli kesim frekansları arasındaki frekans bileşenlerini geçirip bu aralık dışındaki frekans bileşenlerini bastırmaktadırlar. Band geçiren filtreler, filtre kazancının ve çıkış geriliminin maksimum olduğu frekans olan merkez frekansının (tepe noktası) her iki yanında bulunan yarı güç noktaları arasında kalan belirli frekanslarda çalışırlar. Bu frekans sınırları dışarısında kalan frekanslar ise zayıflatılmaktadırlar.



Şekil 13. Band geçiren filtrenin genlik-frekans tepkisi

Kazanç eğrilerinin karakteristik özelliklerine göre çeşitli filtreler bulunmaktadır. Bu çalışmada band geçiren butterworth filtre kullanılmıştır. Butterworth filtrelerin geçirme ve durdurma bandlarında dalgalanma olmazken fonksiyon ise tek düze azalmaktadır. Bu filtrelerin diğer filtrelerden farklı olarak dereceleri arttırıldığında durma bandındaki sert düşme haricinde frekans genlik eğrisinde bir değişme olmaz. Butterworth filtre için tipik bir frekans cevabı şekil (14)'deki gibidir. Bu filtrenin zayıflaması ise eşitlik (2) ile tanımlanır. Eşitlikteki  $w$  istenen zayıflamanın gerçekleştiği frekans,  $w_c$  filtre kesim frekansı,  $n$  filtre eleman sayısını ifade etmektedir.

$$A_{dB} = 10 \log \left[ 1 + \left( \frac{w}{w_c} \right)^{2n} \right] \quad (2)$$



Şekil 14. Butterworth filtre cevabı

Bu çalışmada MATLAB programı aracılığı ile oluşturulan filtrenin kazanç yanıtı ve genlik karakteristikleri incelenerek en uygun filtre derecesi seçilmiş ve istenilen özelliklerde band geçiren filtre tasarımı sağlanmıştır.

### 2.2.2. Öz Nitelik Çıkarma

Bu çalışmada EEG sinyallerine ait verilerden öz nitelik çıkarmak için parametrik yöntemlerden yararlanılmıştır. Parametrik model tekniklerinin amacı bir sistem ya da sinyalin matematiksel model parametrelerini ortaya koymaktır. Bu tekniklerde, bilinen bilgiler modeli tanımlamak için kullanılmaktadır. Parametrik modellemede sistemin fiziksel modeli belirli değildir fakat sistemin temsil edildiği çıkış sinyaline matematiksel bir model uydurulup bu modelin parametreleri belirlenebilmektedir.

EEG işaretleri bir zaman serisi olarak düşünüldüğünde bu işaretlerde zaman içerisinde meydana gelen değişiklikler ölçülüp modellenmektedir. EEG işaretlerinin modellenmesi ve oluşturulan model parametrelerinin kestirimi için çeşitli metotlar bulunmaktadır.

Zaman serileri analizi olarak da adlandırılan parametrik modelleme tekniklerinin başlıcaları Box-Jenkins modelleri olarak bilinen ve durağan olduğu kabul edilen işaretlere uygulanan AR Model Yürüyen Ortalamalı Model (Moving Average model-MA) ve bunların birleşimi olan ARMA modeldir. Modelleme süreci parametrelerle temsil edildiğinden ötürü otoregresif modelleme EEG sinyallerinin analizlerinde de kullanılabilir.

Parametrik yöntemler kullanılırken kullanacak olduğumuz işaretin yapısal özelliklerini de göz önüne almamız gereklidir. EEG gibi anlık oluşan keskin tepelerin bulunduğu işaretler için AR model, bu tür tepelerin olmadığı işaretler için MA model daha uygun iken ARMA model ise her iki tip işaret için de kullanılabilir parametrik yöntemlerdendir. İşlem yükü bakımından ise AR model daha avantajlı olan yöntemdir [26].

Bu çalışmada EEG sinyalleri bir zaman serisi olarak düşünülmüş ve otoregresif süreç sonucunda elde edilen modelden tahminlenen model katsayıları öz nitelik vektörü olarak kullanılmıştır.



### 2.2.2.1. Zaman Serileri Analizi

Zaman serileri bir dönemden diğer döneme değişken değerlerin ardışıl olarak gözlemlenebildiği sayısal büyüklüklerdir. Gözlemlenen verilerin zamanla ardışıl şekilde gerçekleşmesi bir koşul olmayıp düzenli aralıklarla dizinin gelişimini görmek bakımından gerekliliktir. Zaman serilerinin analizleri serinin özelliklerini özetler ve yapısını ortaya koymaya çalışır. Bu işlem hem zaman hem de frekans boyutunda yapılabilmektedir [27]. Zaman serisi analizlerinde, zaman serisi değişkenlerinin gelecek değerlerinin kestiriminde değişkenin geçmiş değerlerinin sahip olduğu bilgiden yararlanılabilmektedir. Zaman serileri için model oluşturulurken, seriyi meydana getiren stokastik sürecin zamanla değişmediği kabul edilir. Burada amaç öngörü yapmak olduğu için zaman serisinin geçmişte göstermiş olduğu değişimlerin iyi bir şekilde incelenmesi gerekmektedir. Öngörü için belirlenecek olan algoritma ise gözlenen zaman serisinin davranışının incelenmesi ile belirlenecektir [28].

Zaman serileri analizi olarak da adlandırılan parametrik modelleme sisteminde, sistem o andaki ve geçmişteki giriş değerleri ile geçmişteki çıkış değerlerinin kombinasyonu ile modellenmektedir. Bu durum denklem (3) ile ifade edilmiştir.

$$y_k = - \sum_{i=1}^p a_i y_{k-i} + \sum_{j=0}^q b_j x_{k-j} \quad (3)$$

Bu modelleme ifadesinde  $x_k$  ve  $y_k$  sistemin ayrık zamandaki giriş ve çıkış serileridir.  $a$  ve  $b$  ifadeleri sistem parametrelerini ifade etmektedir. Model girişi beyaz gürültü olarak alınmakta ve parametre değerleri bilindiği müddetçe geçmiş giriş-çıkış değerlerinden  $k$  anındaki  $y_k$  değeri bulunabilmektedir. Sistem sadece giriş işaretlerinin kullanıldığı MA model olarak da modellenabilir. Bu durumda  $a_i=0$  olur. Bu durum denklem (4) ile gösterilir [3].

$$y_k = \sum_{j=0}^q b_j x_{k-j} \quad (4)$$

Sistem sadece çıkış kullanılarak modellendiğinde ise AR model kullanılmakta olup bu durum  $b_j$  katsayılarının sıfıra eşit olduğu durum ile denklem (5)'deki gibi ifade edilmektedir.

$$y_k = - \sum_{i=1}^p a_i y_{k-i} + b_0 x_k \quad (5)$$

Zaman serilerinde parametre kestirimi için farklı yöntemler geliştirilmiştir, bu yöntemlerin çoğu algoritmalarında En Küçük Kareler yaklaşımını kullanmaktadırlar. Bu teknikte sistem parametrelerini elde etmek için kestirim hatasının karesinin beklenen değerini en küçük yapan ifade kullanılmaktadır.

Zaman serilerinin özelliklerini belirlemek ve gelecek değerlerini öngörebilmek amacını taşıyan Box-Jenkins modellerinde esas amaç incelenen zaman değişkeninin o anki değerini, geçmişteki değerlerinin ağırlıklı toplamı şeklinde ortaya koyabilmektir.

### 2.2.2.2. Box-Jenkins Modelleri

Zaman serileri analizinde doğrusal stokastik süreçlere dayanan Box-Jenkins kestirim yöntemi en yaygın olarak kullanılan istatistiksel yöntemlerdendir. Bu yöntemde bir zaman serisinin modellenebilmesi için serinin durağan olması gerekmektedir. Zaman serisinin durağanlık şartı, serinin ortalama ve varyansında sistematik bir değişme olmamasıdır [29]. Box-Jenkins modelleri durağan zaman serileri için uygulanan üç modelden oluşur. Bunlar AR, MA ve ARMA stokastik süreç modelleridir.

AR model ilk olarak Yule tarafından 1926 yılında önerilmiş ve 1931 yılında Walker tarafından geliştirilmiştir. MA model ise ilk olarak 1937 yılında Slutzky tarafından kullanılmıştır [30].

Bu çalışmada öznitelik vektörü elde etmek amacı ile kullanılan yöntem AR modelleme yöntemidir. AR modelleme sürecinde öznitelik olarak kullanılacak AR katsayılarının elde edilmesinde ise ilerleyen bölümlerde bahsedilecek olan Burg algoritması kullanılmıştır.

#### 2.2.2.2.1. Box-Jenkins Modelinin Kurulumu

Box-Jenkins kestirim yöntemlerinde hedeflenen şey en az parametre ile zaman serisine en uyumlu olan doğrusal stokastik süreç modelini belirlemektir

Box-Jenkins yaklaşımında model kurma dört basamaktan oluşmaktadır. İlk aşamada model belirlenir, ikinci aşamada model tahmini yapılır, bu aşamada model parametrelerinin

kestirimi gerçekleştirilir. Bu işlem için hata karelerinin toplamını minimum yapan parametre değerleri belirlenir. Üçüncü aşama tahmin sonrası kontrol aşamasıdır. Bu bölümde parametreleri tahminlenen modelin verilere uygunluğunun testi gerçekleştirilir. Son aşama ise ileriye yönelik tahmin veya öngörü aşamasıdır. Burada ise zaman serisinin gelecek değerleri tahminlenir [28].

Zaman serilerini meydana getiren sürecin kesin tanımını yapmak için sürecin olasılık dağılımının tanımlanması gereklidir ancak genellikle bu mümkün olmamaktadır. Bu durumda sürecin özelliklerinin saptanabilmesi için otokovaryans fonksiyonu, otokorelasyon fonksiyonu ve kısmi otokorelasyon fonksiyonu kullanılmakta ve bu araçlar süreci tanımlamaya yardımcı olmaktadır [28].

#### 2.2.2.2.2. Otokovaryans ve Otokorelasyon Fonksiyonları

EEG işaretleri gibi değeri tam olarak bilinmeyen yani deterministik olmayan rastgele işaretlerin belirlenebilmesi ve analizlerinin yapılabilmesi için bazı istatistiksel parametrelere ihtiyaç duyulmaktadır. Bu parametreler olasılık yoğunluk fonksiyonu ( $p(x)$ ), varyans ( $\sigma^2$ ), ortalama (beklenen değer, birinci moment- $m$ ) ve karesel beklendik değer (ikinci moment)'dir. Örneğin rastgele bir işaret olan  $x$  için bu parametrelerin matematiksel olarak gösterilişleri sırasıyla aşağıda sunulmuştur;

$$p(x) \quad (6)$$

$$\sigma_x^2 = E((x - m_x)^2) = \int_{-\infty}^{+\infty} (x - m_x)^2 \cdot p(x) \cdot dx \quad (7)$$

$$m_x = E(x) = \int_{-\infty}^{+\infty} x \cdot p(x) \cdot dx \quad (8)$$

$$E(x^2) = \int_{-\infty}^{+\infty} x^2 \cdot p(x) \cdot dx \quad (9)$$

Ayrık zamanlı rastgele değişkenler kullanıldığında yukarıdaki eşitliklerindeki integral ifadelerinin yerini toplama işleci alır.

İki rastgele deęişken olan  $x, y$  arasındaki kovaryansın ifadesi (10) ile gösterilmektedir. Ayrıca bu deęişkenler arasındaki ilişkinin ya da karşılıklı bağımlılığın bir ifadesi olan çapraz ilişki (çapraz korelasyon) (11) eşitliği ile gösterilmektedir.

$$\text{cov}(x, y) = E[(x - E(x))(y - E(y))] \quad (10)$$

$$r_{xy} = E(xy) \quad (11)$$

İki deęişken arasındaki doğrusal ilişkinin derecesi ve yönü hakkında bize bilgi veren korelasyon katsayısı ( $\rho_{xy}$ ) ise (12) ile ifade edilmektedir.

$$\rho_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} \quad (12)$$

Bir zaman serisi işaretinin durağan kabul edilebilmesi için istatistiksel parametrelerinin zamanla deęişmemesi gerekmektedir. İstatistiksel parametrelerden olan korelasyon fonksiyonu zamanın fonksiyonu olmamasına rağmen birbirleriyle ilişkiye sokulan  $x$  ve  $y$  işaret parçaları arasındaki  $\tau$  gecikmenin fonksiyonudur. Bu sebeple  $x$  ve  $y$  zamanla deęişen, sürekli durağan işaretler olduğunda çapraz ilişki fonksiyonu eşitlik (13) gibi olur.  $y(t) = x(t)$  olduğu durumlarda ise beklenen deęer, Oto Korelasyon (öz ilişki) ismini almaktadır. Bu durumun matematiksel ifadesi (14) ile gösterilmiştir. (13) ve (14)'deki eşitliklerin ayrık zamanlı işlemler için ifade edilişi ise sırasıyla (15) ve (16) ile gösterilmiştir [3].

$$r_{xy}(\tau) = E[x(t) \cdot y(t - \tau)] \quad (13)$$

$$r_x(\tau) = r_x(-\tau) = E[x(t) \cdot x(t - \tau)] \quad (14)$$

$$r_{xy}(n) = E[x(i) \cdot y(i - n)] \quad (15)$$

$$r_x(n) = r_x(-n) = E[x(i) \cdot x(i - n)] \quad (16)$$

### 2.2.2.2.3. Otoregresif (Auto Regressive-AR) Model

Yaygın bir kullanıma sahip olan otoregresif model EEG gibi biyolojik sinyallerden nitelik çıkarmak için kullanılan doğrusal özellik çıkarma yöntemlerinden birisidir. AR modelleme, sinyali modellemek için bir veri noktasının kendisinden önceki birkaç veri noktasıyla ilişkili olduğunu varsayar. Bu yaklaşım biyolojik sinyallerin modellenmesi için uygun bir yaklaşımdır.

Otoregresif yöntemde sinyale ait belirli bir andaki genlik değeri sinyalin o zamana kadar örneklenmiş bütün bölümlerinin genliklerinin farklı oranlarda toplanıp sonuca bir tahmin hatası ilave edilmesi ile bulunur. Otoregresif model, sinyal spektrumlarında aniden oluşan tepeliklerin belirlenmesinde kullanılabilir olması ve model parametrelerinin çözümlenmesinde doğrusal denklemlerden yararlandığı için geniş bir kullanıma sahiptir.

AR model çoğunlukla durağan zaman serileri için kullanılır. Belirli bir zamandaki gerçel değerli, sıfır ortalamaya sahip, durağan, deterministik olmayan ve p'inci dereceden AR model denklem (17) ile tanımlanır [31].

$$y_k = -\sum_{i=1}^p a_i y_{k-i} + e_k \quad ; \quad k = 1, 2, \dots, N \quad (17)$$

Denklem (17)'de p, modelin derecesini,  $y_k$  k noktasındaki işaretin değerini,  $a_i$ , gerçel değerli AR katsayılarını, N örnek sayısını,  $e_k$  ise beyaz gürültü hata terimini ifade etmektedir.

AR katsayıları farklı metotlar kullanılarak hesaplanabilmektedir. Burg ve Levinson-Durbin algoritmaları bu metotlar arasındadır. Levinson Durbin algoritmasında öz ilişki fonksiyonlarından yararlanılarak Yule-Walker denklemleri çözümlenir ve AR katsayıları elde edilir.

#### 2.2.2.2.3.1. Yule-Walker Yöntemi

AR modelleme sinyalden uygun bir denklem elde etmek için gerekli olan bir süreç olarak kabul edilebilir. k anından önceki geçmiş çıkış değerleri kullanılarak k anındaki işaret denklem (18)'deki şekilde kestirilip kestirim sonucunda oluşan hata ise denklem (19)'daki gibi ifade edilir. Denklem (19) aynı zamanda bilinmeyen girişin kestirim ifadesidir [3].

$$\hat{y}_k = -\sum_{i=1}^p \hat{a}_i y_{k-i} \quad (18)$$

$$e_k = y_k - \hat{y}_k = y_k + \sum_{i=1}^p \hat{a}_i y_{k-i} \quad (19)$$

Denklem (19)'da ifade edilen hata teriminin en küçük olabilmesi için denklem (20) ve denklem (21)'deki ifadeler sağlanabilmelidir.

$$\frac{dE(e_k^2)}{d\hat{a}_i} = 0 \quad (20)$$

$$2E((y_k + \sum_{i=1}^p \hat{a}_i y_{k-i}) y_{k-i}) = 0 \quad (21)$$

Denklem (21)'in daha önce tanımlanan özilişki fonksiyonu yardımıyla çözümlenmesinden p adet ve p bilinmeyene sahip doğrusal denklem takımı elde edilir. Bu denklem takımı (22) ile gösterilmiştir.

$$\begin{bmatrix} r_1 & + \hat{a}_1 r_0 + \hat{a}_2 r_1 + \dots & + \hat{a}_p r_{p-1} \\ & \vdots & \vdots \\ r_p & + \hat{a}_1 r_{p-1} + \hat{a}_2 r_{p-2} + \dots & + \hat{a}_p r_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (22)$$

(22)'deki eşitlikler Yule-Walker eşitlikleri olarak bilinmektedir. Buradaki  $r_0, r_1, r_2, \dots, r_p$  toplam p+1 adet özilişki ifadesi belirlendiği takdirde,  $a_1, a_2, \dots, a_p$  olmak üzere p tane en iyi parametre değeri belirlenebilir. (22)'deki Yule-Walker eşitliklerini matris şeklinde ifade edecek olursak denklem (23) elde edilir. Denklem (23)'de ki R, simetrik ve köşegenleri üzerinde bulunan elemanları aynı olan Toeplitz özilişki matrisidir. Bu matrisin tersi alınarak denklem (24)'de ki AR katsayılarının oluşturduğu vektör elde edilir [3].

$$R \cdot \vec{a} = -\vec{r} \quad (23)$$

$$\vec{a} = -R^{-1} \cdot \vec{r} \quad (24)$$

Denklem (24)'de görüldüğü gibi özilişki matrisinin tersini alma işlemi yerine Kısmi Korelasyon yöntemi olarak isimlendirilen bir yöntem Levinson Durbin tarafından geliştirilmiştir. Bu yöntem (25-29) eşitlikleri ile elde edilir. Bu eşitlikler  $i=1, 2, 3, \dots, p$  için sırayla çözümlenir ve çözümlene sonucunda  $p$ . Dereceden AR katsayıları (28) eşitliğindeki gibi elde edilir [3].

$$E_0 = r_0 \quad (25)$$

$$k_i = -\frac{r_i + \sum_{j=1}^{i-1} a_j^{(i-1)} \cdot r_{i-j}}{E_{i-1}} \quad (26)$$

$$a_i^1 = k_i; \quad i = 1, 2, 3, \dots, p \quad (27)$$

$$a_j^i = a_j^{i-1} + k_i a_{i-j}^{i-1}; \quad j = 1, 2, 3, \dots, (i-1) \quad (28)$$

$$E_i = (1 - k_i^2) \cdot E_{i-1} \quad (29)$$

$$a_i = a_i^p \quad (30)$$

Eşitliklerdeki  $k_i$  katsayıları, Kısmi Korelasyon ya da Yansıma katsayıları olarak isimlendirilir. Yule-Walker eşitliklerinin Levinson Durbin yöntemi olan Kısmi Korelasyon yöntemi ile ardışıl olarak çözümlenmesi hesaplama işlemlerinin daha hızlı ve kolay gerçekleşmesini sağlamaktadır. Fakat model derecesinin kararlı bir şekilde artırılabilmesi için  $E_i$  kazancının da azalması gerekmektedir. Bu sebepten ötürü (30) eşitliği incelendiğinde  $k_i$  katsayılarının 1'den küçük olma gerekliliği görülmektedir. Levinson Durbin algoritması ise bu durumu garanti etmemektedir [3].

#### 2.2.2.2.3.2. Burg Yöntemi

Burg yinelemeli algoritması daha çok veri kullanarak ileri-geri hata minimizasyonu yapmaktadır. Bu yönteme aynı zamanda Maksimum Entropi ismi de verilmektedir. İleri ve geri ortalama karesel tahmin hataları denklem (31) ve denklem (32) ile gösterilmektedir.

Burg yönteminde bu hataların toplamını en aza indirmeye çalışılacak şekilde iterasyon yapılır [3].

$$f_p(n) = \sum_{k=0}^p a_k^p \cdot y(n-k) \quad (31)$$

$$b_p(n) = \sum_{k=0}^p a_k^p \cdot y(n-p+k) \quad (32)$$

Denklem (33) ve denklem (34) ise ileri ve geri tahmin hataları arasındaki ilişkiyi, (35) ise ortalama karesel hata toplamlarını göstermektedir.

$$f_p(n) = f_{p-1}(n) - \gamma_p \cdot b_{p-1}(n-1) \quad (33)$$

$$b_p(n) = b_{p-1}(n-1) - \gamma_p \cdot f_{p-1}(n) \quad (34)$$

$$Q_p = E(f_p^2(n)) + E(b_p^2(n)) \quad (35)$$

Burg yönteminde amaç denklem (35)'de gösterilen hatayı minimum yapmaktır. Yukarıdaki eşitliklerdeki  $\gamma_p$  Yansıma Katsayısı olarak adlandırılmaktadır.  $Q_p$  hatasını en küçük yapacak  $\gamma_p$  değerini belirleyebilmek için (36)'deki gibi hatanın yansıma katsayısına göre türevi alınıp sıfıra eşitlenir sonrasında ise (33), (34) ve (35)'nin ortak çözümünden (37) ve (38) eşitliği elde edilir [3].

$$\frac{dQ_p}{d\gamma_p} = 2E\left(f_p(n) \frac{df_p(n)}{d\gamma_p} + b_p(n) \frac{db_p(n)}{d\gamma_p}\right) = 0 \quad (36)$$

$$E(b_{p-1}(n-1) (f_{p-1}(n) - \gamma_p b_{p-1}(n-1)) + f_{p-1}(n-1) (b_{p-1}(n-1) - \gamma_p f_{p-1}(n))) = 0 \quad (37)$$

$$\gamma_p = \frac{2 \cdot \sum_{n=p}^{N-1} f_{p-1}(n) \cdot b_{p-1}(n-1)}{\sum_{n=p}^{N-1} (f_{p-1}^2(n) + b_{p-1}^2(n-1))} \quad (38)$$



(38) ile elde edilen yansıma katsayısının değeri 1'den küçük olup kararlılık kriterini sağlayacaktır. Sonrasında ise (39) da gösterilen matrissel ifadedeki gibi Levinson Durbin yöntemindeki benzer bir iterasyon ile p-1 dereceli kestirim filtresine ait katsayılarından p dereceli kestirim filtresine ait katsayılara geçiş gerçekleştirilir. (40) ile (41) ise yönteme ait başlangıç koşullarıdır [3].

$$\begin{vmatrix} 1 \\ a^p(1) \\ a^p(2) \\ \cdot \\ \cdot \\ a^p(p-1) \\ a^p(p) \end{vmatrix} = \begin{vmatrix} 1 \\ a^{p-1}(1) \\ a^{p-1}(2) \\ \cdot \\ \cdot \\ a^{p-1}(p-1) \\ 0 \end{vmatrix} - \gamma_p \cdot \begin{vmatrix} 0 \\ a^{p-1}(p-1) \\ a^{p-1}(p-2) \\ \cdot \\ \cdot \\ a^{p-1}(1) \\ 1 \end{vmatrix} \quad (39)$$

$$f_0(n) = b_0(n) = y(n) ; \quad 0 \leq n \leq (N-1) \quad (40)$$

$$E_0 = \frac{1}{N} \sum_{n=0}^{N-1} y^2(n) \quad (41)$$

Ortalama kestirim hatası ise (42) ile gösterilir.

$$E_p = E_{p-1} \cdot (1 - \gamma_p^2) \quad (42)$$

Burg yöntemi işlemsel açıdan verimli bir yöntem olmakla birlikte kararlı bir otoregresif yöntem sağlar.

Performansı yüksek ve kararlı bir AR modeli elde edebilmek için bir takım hususlara dikkat edilmesi gerekmektedir. Bu hususlar;

- Uygun algoritmanın belirlenmesi
- Uygun model derecesinin seçimi
- AR modelin elde edilecek olduğu işaretin boyutu
- Verinin durağanlık derecesidir

AR model derecesinin belirlenmesi önemli bir husustur. Model derecesi modeldeki parametre sayısını belirler, belirlenen derece büyük seçilirse spektrumda hatalı ve yanıltıcı tepcikler oluşacağından spektrum bozulmaları meydana gelir. Derece küçük seçildiğinde

ise spektrumda belirli tepecikler oluşamayacağı için spektrumun frekans içeriği tam olarak belirlenemez [32].

AR model derecesinin belirlenebilmesi için literatürde birçok yöntem önerilmiştir. Bunlar arasından en yaygın kullanıma sahip olan yöntem Akaike tarafından önerilen Akaike Bilgi Kriteri (Akaike Information Criterion-AIC) fonksiyonudur. AIC'e göre ideal model derecesi (43) eşitliğini minimum yapan derecedir. Bu eşitlikteki  $p$  model derecesini,  $\sigma_p^2$   $p$ . derecedeki hata dizisinin varyansı,  $N$  ise veri uzunluğunu ifade etmektedir.

$$AIC(p) = \ln(\sigma_p^2) + \frac{2p}{N} \quad (43)$$

### 2.2.3. Tahmin ve Etme ve Sınıflandırma

#### 2.2.3.1. Kısmi En Küçük Kareler Regresyon Yöntemi

KEKKR günümüzde pek çok teknolojik ve bilimsel çalışmalarda kullanılan çok değişkenli istatistiksel bir yöntem olup kısmi en küçük kareler analizi (KEKK) ve çoklu doğrusal regresyon analizinin birleşiminden meydana gelmektedir. Bu yöntem özellikle çoklu doğrusal regresyon çalışmalarında, fazla sayıdaki bağımsız değişkenler arasında bir ya da daha fazla doğrusal bağlantının olduğu ve çoklu bağlantı olarak adlandırılan sorunun ortaya çıktığı durumlarda kullanılmaktadır. KEKKR ile bağımlı ve bağımsız bileşenlerdeki değişimi yüksek oranda açıklayan, aralarında çoklu doğrusal bağlantı sorunu olmayan, kestirim bakımından önceki duruma göre daha az bileşen sayısına sahip modeller oluşturulabilmektedir.

KEKKR ile ilgili ilk çalışmalar 1970'li yılların başlarında Herman Wold tarafından ekonometri alanında yapılmıştır [33]. Bu bakımdan KEKKR yöntemi sosyal bilimlerin kökenlidir. Daha sonra oğlu olan Svante ise bu yöntemi kemometri alanında geliştirmiştir. Yöntem ile ilgili günümüze değin birçok çalışma yapılmıştır. Bunlar arasında en kapsamlı olanı Geladi ve Kowalski tarafından 1986'da yapılanıdır [34]. Höskuldsson ise KEKKR yöntemini istatistiksel açıdan incelemiştir [35]. İstatistiksel açıdan değerlendirme çalışmaları Phatak ve de Jong ile Braak ve de Jong tarafından devam ettirilmiştir [36-37]. Son dönemlerde bu yöntemle ilgili yapılan çalışmalara Wold ve arkadaşları, Tobias, Abdi ve Polat'ın yapmış oldukları çalışmalar örnek olarak gösterilebilir [38-40].

### 2.2.3.1.1. Çoklu Doğrusal Regresyon

İki ya da daha çok değişken arasında bir ilişkinin olup olmaması ve eğer bir ilişki var ise bunun derecesinin belirlenmesi istatistiksel çözümlerle sıklıkla karşılaşılan bir problemdir. Regresyon analizi de iki ya da daha çok değişken arasındaki ilişkiyi ölçmek için kullanılan bir metottür. Bu analiz yöntemi ile değişkenler arasındaki ilişki ve bu ilişkinin gücü hakkında bilgi edinilir. Regresyon analizinde değişkenler bağımlı ve bağımsız değişkenler olmak üzere ikiye ayrılmaktadır. Bağımlı değişkenler, bağımsız olan değişkenler tarafından açıklanmaya çalışılırlar. Regresyon analizi bağımsız değişken sayısına göre; basit regresyon analizi ve çoklu regresyon analizi, fonksiyon tipine göre ise doğrusal regresyon analizi ve doğrusal olmayan regresyon analizi olmak üzere alt gruplara ayrılmaktadır. Birden fazla bağımsız değişkenle yapılan analize çoklu regresyon analizi denilmektedir.

Herhangi bir bağımlı değişkenin bir veya daha fazla bağımsız (açıklayıcı) değişken ile arasındaki ilişkinin matematiksel ifadesi olan klasik doğrusal regresyon modeli (44) ile ifade edilebilir.

$$y = 1\beta_0 + X\beta + \varepsilon \quad (44)$$

Burada amacımız regresyon denklemi aracılığı ile açıklayıcı değişkenlerin çeşitli değerlerine denk gelen bağımlı değişkenlerin değerinin tahmin edilmesidir.

(44) eşitliğinde  $n$  gözlem sayısı,  $p$  bağımsız değişken sayısı olmak üzere  $y$ , ( $n \times (p + 1)$ ) boyutlu bağımlı değişkenlerden oluşan gözlem vektörünü,  $X$ , ( $n \times p$ ) boyutlu bağımsız değişkenlerden oluşan gözlem vektörünü,  $\varepsilon$  ise ( $n \times 1$ ) boyutlu hata vektörünü ifade etmektedir.

(44) ile ifade edilen modelde  $X$  gözlem vektöründeki bilinmeyen  $\beta_0$  sabitine denk gelen 1'ler den oluşan ilk sütun merkezileştirme yöntemi ile silinebilmektedir. Bu amaçla açıklayıcı değişkenlerin her bir tanesi için toplam  $n$  tane gözlem değerinden elde edilen sütun ortalamaları ilgili olan değişkenlerden çıkartılır. Merkezileştirme sonucunda  $X$  matrisindeki birlerden meydana gelen ilk sütun,  $\beta$  vektöründen ise baştaki  $\beta_0$  silinir. Sonuç olarak (44)'deki model (45)'deki hale dönüşür [40].

$$y = X\beta + \varepsilon \quad (45)$$

Modelin yeni durumuna göre  $y$ ,  $(nx1)$  boyutlu bağımlı değişkenler için gözlem vektörü,  $X, (nxp)$  boyutlu açıklayıcı değişkenlerden oluşan gözlem vektörü,  $\varepsilon$   $(nx1)$  boyutlu hata vektörü,  $\beta$  ise  $(px1)$  boyutlu bilinmeyen doğrusal regresyon katsayıları vektörüdür. Aynı model birden fazla bağımlı değişken kullanılarak yapıldığında oluşan doğrusal regresyon modelinin matrisel gösterimi ise (46)'daki gibidir. Bu ifadedeki bağımlı değişkenlerden oluşan  $Y$  gözlem vektörü  $(nxm)$  boyutlu, bağımsız değişkenlerden oluşan  $X$  gözlem vektörü  $(nxp)$  boyutlu, hata vektörü olan  $E$ ,  $(nxm)$  boyutlu ve regresyon katsayılarından oluşan  $B$  vektörü ise  $(pxm)$  boyutludur. Modelde hataların birbirleriyle ilişkisiz, aynı varyansa sahip ve normal dağılımlı oldukları kabul edilir [36].

$$Y = XB + E \quad (46)$$

### 2.2.3.1.1.1. En Küçük Kareler Yöntemi

En küçük kareler (EKK) yöntemi çok sayıdaki değişken arasındaki ilişkiyi ortaya koyan doğrusal regresyon denkleminin kestirimi için kullanılan bir yöntemdir. Bu yöntem bağımlı ve bağımsız değişkenler arasındaki ilişkiyi en iyi şekilde kestirecek olan parametreleri hesaplar. Güvenilir bir regresyon analizinde hedeflenen şey tahminlenen değer ile gerçek gözlem değeri arasında farkın olmaması ya da bu farkın olabildiğince küçük olmasıdır. Bu amaçla EKK yöntemi parametre hesabını yaparken artık kareler toplamının en küçük değer almasına dayanan yöntemi kullanır.

Hata terimi (47) ile gösterildiği gibi her bir gözlem vektöründeki gerçek değer ile modelden elde edilen tahmin değeri arasındaki farktır.

$$\varepsilon = y - X\hat{\beta} \quad (47)$$

$$\sum e_i^2 = \sum (y_i - \hat{y}_i)^2 = 0 \quad (48)$$

$$\frac{d((y-X\hat{\beta})'(y-X\hat{\beta}))}{d\hat{\beta}} = 0 \quad (49)$$

$$\hat{\beta} = (X'X)^{-1}X'y \quad (50)$$

Klasik EKK tahmin edicisi olan  $\hat{\beta}$ 'nin değeri (49) eşitliğinin çözümlenmesi ile elde edilir. Bu eşitlikteki  $(X'X)^{-1}$  ifadesinin elde edilebilmesi için  $X$  sütunlarının doğrusal bağımsız olması gerekmektedir. Eşitlikteki  $X$  matrisi  $(p+1)$  tam ranka sahip bir matris ise  $\hat{\beta}$  tahmin edici (50) eşitliği ile elde edilebilir. Klasik EKK yöntemi ile regresyon katsayıları kestiriminin yapılabilmesi için  $(X'X)^{-1}$  matris rank değerinin  $(p+1)$ 'e eşit olması ve değişken sayısı olan  $p$  değerinin  $n$  gözlem sayısına eşit ya da gözlem sayısından küçük olması gerekmektedir ( $p \leq n$ ). [34, 41]. Çok sayıda açıklayıcı değişken kullanılması durumunda (50)'deki tahminleyici ifadesi (51)'deki şekilde gösterilecektir.

$$\hat{B} = (X'X)^{-1}X'Y \quad (51)$$

### 2.2.3.1.1.2. Çoklu Bağlantı

Açıklayıcı değişken sayısının fazla olduğu durumlarda değişkenler arasında çoklu doğrusal bağlantı sorunu ortaya çıkabilmektedir. Bu sorun analizler sonucunda elde edilen EKK tahminleyicilerinin varyans değerlerinin çok büyük olup kestirimlerin asıl değerlerinden uzaklaşmasına ve hatalı tahminlere sebep olmaktadır [46] Ayrıca gözlem sayısından daha fazla sayıda değişkenin olması durumu da ( $n < p + 1$ ) klasik EKK regresyonunu kullanılamaz hale getirmektedir [42].

Çoklu bağlantı durumu tam çoklu bağlantı ve güçlü çoklu bağlantı durumu olmak üzere ikiye ayrılmaktadır. Tam çoklu bağlantı durumu açıklayıcı değişkenler arasındaki ilişkinin tamamen doğrusal olduğu durumlarda oluşur. Bu durumlarda açıklayıcı değişkenlerden biri diğerleri cinsinden ifade edilebilmekte ve  $X'X$  matrisinin rankının  $(p + 1)$ 'den küçük olacak olmasından ötürü tersi alınamamaktadır. Dolayısıyla  $\beta$ 'nin bir tane klasik EKK tahmini olmamaktadır. Açıklayıcı değişkenler arasında tam bir bağımsızlık durumunun olmadığı durumlarda ise güçlü çoklu bağlantı durumu ortaya çıkmaktadır. Bu tip çoklu bağlantılarda yaklaşık olarak sıfıra eşit olan doğrusal bağlantılar söz konusudur ve  $X'X$  matrisinin tersi alınabilir olup matematiksel olarak tahmin edici tektir. Fakat güçlü çoklu bağlantı regresyon sonuçlarında belirsizliğe neden olup başarısız tahminler oluşturmaktadır [43-45].

Bu gibi durumlarda bu sorunlara çözüm getiren KEKKR kullanılabilir. KEKKR yöntemine geçmeden önce bu yöntemin temelini oluşturan KEKK yönteminden bahsedilecektir.

### 2.2.3.1.2. Ayarlama

Yukarıda bahsedilen çoklu bağlantı problemi aynı zamanda çok değişkenli ayarlama için istatistiksel açıdan sorun teşkil etmektedir. Bu sorunun çözülmesi için önemsiz değişkenlerin modelden çıkarılarak değişken sayısının azaltılması gereklidir. Azaltma işlemi ise çoğu zaman daha az doğrulukta tahminler yapılmasına sebebiyet vermektedir. Bu nedenle çoklu doğrusal regresyondaki değişken sayısını azaltma işlemi yerine ileride bahsedilecek olan KEKKR yöntemi kullanılarak açıklayıcı  $X$  değişkenlerinin daha az sayıdaki indirgenmiş doğrusal birleşimi üzerinden  $Y$ 'yi döndürerek, yani ters ayarlama işlemi ile yeni değişkenler elde edilir ve hesaplamalarda bu değişkenler kullanılır.

Ayarlama, veri kümesindeki  $Y$  gözlemlerinden bir takım matematiksel dönüşümler kullanılarak bilinmeyen  $X$  bilgisinin elde edilebilmesi için var olan bilgilerin çeşitli yöntemlerle kullanılmasıdır. Buradaki amaç rastlantı değişkeni olmayan  $X$  için ileriye dönük kestirimler yapmaktır. Ayarlama probleminde bağımlı değişkenlerin tahminlenmesi klasik veya ters kestirim metotlarıyla incelenir [47].

Ayarlamada bütün değişkenler için kullanılacak değerler merkezileştirilip ölçeklendirilir. Merkezileştirme işleminde her bir değişken için ortalama değer hesaplanıp ilgili değişkenden çıkartılır. Ölçeklendirme işlemi için ise ayarlama kümesindeki değişkene ait bütün değerler o değişken için standart sapmaya bölünür [34]. Tek değişkenli ayarlama olduğu gibi çok değişkenli ayarlama da amaçlanan şey verilen  $y$  rastlantı değişkenlerinden  $X$  açıklayıcı değişkenlerinin ön kestirimidir [47]. Yani  $\hat{Y} = f(X)$  olacak şekilde model kurulup modelin parametreleri kestirilir. Bu amaçla kullanılan ayarlama modelleri dört tanedir; ileri regresyon modeli, karışım modeli, genişletilmiş karışım modeli, gizli değişkenler üzerinden regresyon modeli.

İleri regresyon ve karışım modeli sırasıyla (52), (53)'deki gibi ifade edilir. Karışım modelinin genişletilmiş hali olan genişletilmiş karışım modeli ise (54)'deki gibidir. Çoklu doğrusal regresyon modeli ileri regresyon modelinin bir örneğidir ve  $(Y = XB + F)$  şeklindeki modellemeyi kullanarak  $Y$  gözlemlerini  $X$  üzerinden döndürerek elde eder. Bu

modelleme çoklu doğrusal regresyon kestiricisi olan  $\hat{y}_i = x_i'(X'X)^{-1}X'Y$  'i verir. Bu ayarlama türü ters ayarlama olarak da adlandırılmaktadır. Çoklu doğrusal regresyon ve KEKKR ters ayarlama metotlarıdır [kısmiregrestez-H]. Gizli değişkenler üzerinden regresyon model denklemleri ise (55), (56), (57)'de verilmiştir. KEKKR da gizli değişkenler üzerinden regresyon modelleme türüne örnek olarak gösterilebilir.

$$Y = f(X) + F \quad (52)$$

$$X = h(Y) + E \quad (53)$$

$$X = h(Y, T) + E \quad (54)$$

$$T = h_1(X) \quad (55)$$

$$Y = h_2(T) + F \quad (56)$$

$$X = h_3(T) + E \quad (57)$$

Bu modellerdeki E ve F artıkları, T ise tanımlanamayan sistematik yapıyı ifade eder. Ayarlama işlemlerinde Y kullanılarak X'in açıklandığı modeller nedensel, X kullanılarak Y'nin açıklandığı modeller ise kestirici modellerdir. Ayarlama istenilen ise X kullanılarak Y'nin kestiriminin gerçekleştiği ileri yönlü kestirimdir. Bu model yapılarını daha iyi kavrayabilmek için veri sıkıştırma yöntemleri ve bilinear modellemeden bahsedilecektir.

#### 2.2.3.1.2.1. Veri Sıkıştırma

Bir ayarlama model türü olan veri sıkıştırma işleminde çok sayıdaki gözlem verisi, skorlar ya da faktörler olarak da isimlendirilen gizli bileşenler şeklinde (58)'deki gibi sıkıştırılabilir. Bu işlem sonucunda elde edilen bileşenler (59)'daki gibi regresyon denkleminde açıklayıcı değişkenler olarak kullanılırlar.

$$(t_1, t_2, \dots, t_A)' = h_1 [(x_1, x_2, \dots, x_p)'] \quad (58)$$

$$(y_1, y_2, \dots, y_m)' = h_2 [(t_1, t_2, \dots, t_A)'] + f' \quad (59)$$

Bu denklemlerdeki  $A$  kestirim için gerekli olan bileşen sayısını,  $p$  açıklayıcı değişken sayısını,  $m$  bağımlı değişken sayısını,  $f$  ise indirgenmiş bileşenler tarafından açıklanamayan kısmı ifade etmektedir. Yukarıdaki eşitliklerde görüldüğü gibi çok sayıdaki açıklayıcı  $X$  değişkenleri daha az sayıda olan  $T$  değişkenlerine indirgenmekte ve böylece kestirim sonucunda elde edilmesi gereken model parametre sayısı da azalmaktadır. Bu durum çoklu bağlantı sorununun yaşandığı veri kümeleri için çözüm sunmaktadır.  $X$  ve  $Y$ 'nin merkezileştirilmesi sonrasında (60) eşitliği elde edilmektedir. Eşitliklerdeki  $V$  ve  $Q$  değerlerinin elde edilmesinden sonra ise (61)'deki kestirici eşitliği bulunur. Bu işlemlerde dikkat edilmesi gereken nokta ideal bileşen sayısıdır. Çünkü indirgenen bileşen sayısı bağımsız açıklayıcı değişken sayısına eşit olursa yöntem çoklu doğrusal regresyona yakınsar ve çoklu bağlantı sorununu çözüme gücünü kaybeder. Aynı zamanda çok fazla seçilen bileşen sayısı da kestirilecek parametre sayısı artacağından tahminlerin hatalı olmasına sebebiyet verebilir. Bu bakımdan ideal durum bileşen sayısı çoğunlukla açıklayıcı değişken sayısından az seçilmektedir [48].

$$T = XV \quad (60)$$

$$Y = TQ' + F \quad (61)$$

### 2.2.3.1.2.2. Bilineer Yöntemler

Veri sıkıştırma işlemi için çeşitli metotlar mevcuttur. Bu metotlar içerisinde  $X$  ve  $Y$  arasındaki ilişki hakkında çok fazla bilgi gerektirmeyen yöntemlere bilineer yöntemler denilmektedir. KEKKR'da bilineer yöntemdir. Bu yöntemler (60)'deki  $V$  değerini ayarlama verisinden kendileri tahmin ederler [48]. Bilineer yöntemler genellikle diğer veri sıkıştırma yöntemleri gibi ileri ayarlama yöntemleri olarak uygulanmaktadır. Sadece diğerlerinden farklı olarak bilineer yöntemler  $V$ 'e ait elemanların tahminini veri kümesinden kendileri yaparlar. Bilineer modellemede EKK benzeri kestirim yapılır. Bu



modellemede  $X$  kendisine (57)'de ifade edilen bir modelle yakınsar. Bu model  $T$  (skorlar) ve  $P$  (yükler)'den meydana gelmektedir. Merkezileştirilmiş ve ölçeklendirilmiş bilinear ayarlama modeli (63) ve (64)'da ki gibi verilmektedir [40].

$$(T = XV) \quad (62)$$

$$X = TP' + E \quad (63)$$

$$Y = TQ' + F \quad (64)$$

Eşitliklerdeki  $E$  ve  $F$  artıkları,  $P$  ve  $Q$  ise sırasıyla  $X$ 'in ve  $Y$ 'nin  $T$  üzerinden regresyon katsayılarını ifade etmektedir. Bilinear ayarlama yönteminde parametre kestirimi özetlenecek olursa; ilk olarak  $X$  ve  $Y$  için merkezileştirme yapılır. Daha sonra  $\hat{V}$  matrisinin belirlenebilmesi için yöntemi tanımlayacak bir ölçüt optimize edilir.  $\hat{T} = X\hat{V}$  ile skorlar bulunur. Sonrasında ise doğrusal regresyon aracılığıyla  $\hat{P}' = (\hat{T}'\hat{T})^{-1}\hat{T}'X$  ve  $\hat{Q}' = (\hat{T}'\hat{T})^{-1}\hat{T}'Y$  yükleri kestirilir. Artıkları hesaplamak için ise  $\hat{E} = X - \hat{T}\hat{P}'$  ve  $\hat{F} = Y - \hat{T}\hat{Q}'$  kullanılır [48].

Buraya kadar olan kısım ayarlama bölümü idi. Ayarlama ile  $V, T, P, Q, E, F$  parametrelerinin tahminleri yapılır. Sonrasında ise yeni bir  $\hat{y}_i$ 'i kestirmek için merkezileştirilmiş  $X$  vektörü  $\hat{V}$  ile çarpılarak  $\hat{t}'_i = (x'_i - \bar{x}')\hat{V}$ ,  $\hat{t}_i = (\hat{t}_{i1}, \dots, \hat{t}_{iA})$  bileşenleri elde edilir. Buradaki  $A$  sayısı  $X$ 'in indirgenmiş bileşen sayısıdır. Daha sonra bulunan bileşenleri  $\hat{Q}$  ile çarpılıp sonuca  $\bar{y}$  ( $y$ 'nin ortalaması) ilave edilir ve  $\hat{y}'_i = \bar{y}' + \hat{t}'_i\hat{Q}$  kestiricisi elde edilir [48].

### 2.2.3.1.3. Kısmi En Küçük Kareler Yöntemi

Bu yöntem Herman tarafından 1960'lı yıllarda geliştirilmiştir. Kısmi en küçük kareler yöntemi ile aralarında çoklu doğrusal ilişki bulunan açıklayıcı değişkenler çeşitli algoritmalar aracılığıyla hem açıklayıcı hem de bağımlı değişkenlerdeki değişimi açıklayan, aralarındaki çoklu doğrusal bağlantı sorunu ortadan kalkmış ve önceki açıklayıcı değişken sayısından daha az sayıda bileşene indirgenmektedir [42]. Bu yöntemde amaç veri kümesindeki bağımlı ve açıklayıcı değişkenler arasındaki kovaryansı

en çoklayan bileşen sayısını bulmaktır. KEKK yönteminde analiz işlemlerinden önce bütün değişkenlerin birbirleriyle aynı önem sırasında olmaları için merkezileştirme ve sonrasında ise ölçeklendirme işlemi yapılır. Merkezileştirme işlemi daha önceki bölümlerde de bahsedildiği gibi bağımlı ve açıklayıcı tüm değişkenlerden ortalama değerlerinin çıkartılmasıdır. KEKK yönteminde açıklayıcı değişkenlerin boyutunun indirgenmesi ile gizli değişkenler elde edilmektedir. Elde edilen gizli değişkenler açıklayıcı değişkenlerin doğrusal birleşimidirler. Bu değişkenler bağımlı değişkenleri açıklayıcı değişkenler ile döndürmek için kullanılmaktadırlar. Ayrıca gizli değişkenler yeni açıklayıcı değişkenler olarak ileride bahsedilecek olan regresyon analizi için de kullanılmaktadırlar. KEKK'de açıklayıcı değişkenler (65)'deki gibi  $X$ 'in bilineer ayrışımı şeklinde ifade edilirler.

$$X = t_1 p'_1 + t_2 p'_2 + t_3 p'_3 + \dots + t_p p'_p = \sum_{j=1}^p t_j p'_j = TP' \quad (65)$$

Yukarıdaki eşitlikteki  $Xr_j$  şeklinde yazılan  $t_j$ 'ler  $X$ 'in doğrusal birleşimidirler.  $r_j$  ağırlıkları birim dik olmayıp  $t_j$ 'ler ise diktirler.  $p_j$ 'ler yükler olarak isimlendirilip ( $px1$ ) boyutludurlar. Eşitliğin matrissel gösteriminde ise  $T$  ve  $P$  ifadeleri kullanılmıştır. KEKK yönteminde KEKK tahmin edicilerinin ( $\beta$  ve  $B$ ) bulunması için çeşitli algoritmalar mevcuttur. Doğrusal olmayan yinelemeli en küçük kareler (Non Linear Iterative Partial Least Squares-NIPALS) ve KEKK yönteminin istatistiksel olarak esinlenilmiş değişikliğinin basit bir uygulaması (SIMPLS) bu algoritmaların önde gelenleridir. NIPALS algoritmasından literatürde klasik KEKK algoritması olarak bahsedilmektedir. İlerleyen bölümlerde bu algoritmalara ayrıntılarıyla değinilecektir. NIPALS algoritmasında  $t_j$ 'ler (66)'daki gibi  $E_j$ 'lerin doğrusal birleşimlerinden meydana gelirler. Buradaki  $E_j$ 'ler artık matrislerdir. [49].

$$t_j = E_{j-1} w_j \quad (66)$$

$$E_j = X - \sum_{i=1}^j t_i p'_i \quad , E_0 = X \quad (67)$$

Eşitlikteki  $w_j$ 'ler  $r_j$ 'ler ile aynı uzayda bulunan ağırlık kümeleridir ( $j = 1, 2, \dots, h$ ),  $h$  ise matrisin ilgili vektörün ilk  $h$  dizisinden meydana geldiğini belirtir.  $w_j$ 'ler birim diktirler. Birçok algoritmada ilk olarak  $w_j$  ve  $r_j$  değerleri bütün  $j$ 'ler için hesaplatılıp  $t_j$

elde edilir. Sonrasında ise  $p_j$ 'nin hesaplanması için  $X$  matrisi  $t_j$  üzerinde döndürülür.  $h$  tane boyut elde edilmesinden sonra  $T_h$ ,  $P_h$ ,  $R_h$  eşitlikleri bulunur. (70) ağırlık vektörlerinin doğrusal dönüşümle bağlanma ifadesidir. (68) ve (69)'nın ortak çözümünden ise (71) elde edilir.  $h$  adet boyut elde edildikten sonra bağımlı değişkenlerin  $T_h$  üzerine yansımından KEKK ile bulunan  $\hat{y}_{KEKK}^h$  uygun değerler vektörü elde edilir[36, 49].

$$T_h = XR_h \quad (68)$$

$$P_h = X'T_h(T_h'T_h)^{-1} \quad (69)$$

$$R_h = W_h(P_h'W_h)^{-1} \quad (70)$$

$$R_h'P_h = P_h'R_h = R_h'X'T_h(T_h'T_h)^{-1} = T_h'T_h(T_h'T_h)^{-1} = I_n \quad (71)$$

$$\hat{y}_{KEKK}^h = T_h(T_h'T_h)^{-1}T_h'y \quad (72)$$

(72) eşitliğinde gerekli düzenlemeler ile  $y$  yerine  $X\hat{\beta}_{OLS}$  ve  $T_h$  yerine de (68) eşitliği yazılırsa (73) elde edilir.

$$\hat{y}_{KEKK}^h = XR_h(R_h'X'XR_h)^{-1}R_h'X'X\hat{\beta}_{OLS} \quad (73)$$

$$\hat{\beta}_{KEKK}^h = R_h(R_h'X'XR_h)^{-1}R_h'X'X\hat{\beta}_{OLS} \quad (74)$$

(74)'deki KEKK tahmin edicinin daha sade bir görünümü için (68) ve (69)'nın ortak çözümünden elde edilen  $P_h = X'XR_h(R_h'X'XR_h)^{-1}$  (74) eşitliğinde kullanıldığında (75) elde edilir. Bu eşitliğin çok değişkenli gözlemler için karşılığı ise (76) eşitliği ile verilmektedir.

$$\hat{\beta}_{KEKK}^h = R_hP_h'\hat{\beta}_{OLS} = W_h(P_h'W_h)^{-1}P_h'\hat{\beta}_{OLS} \quad (75)$$

$$\hat{B}_{KEKK}^h = R_hP_h'\hat{B}_{OLS} = W_h(P_h'W_h)^{-1}P_h'\hat{B}_{OLS} \quad (76)$$

Eşitlik (76)'daki  $W_h(P_h'W_h)^{-1}P_h'$  matrisi bir izdüşüm matrisidir. Fakat simetrik olmadığından ötürü eğik projektör olarak adlandırılmaktadır. Bu sebepten (75)'deki KEKKR tahmin edicisi  $\hat{\beta}_{OLS}$ 'nin  $W_h$  üzerine eğik izdüşümüdür. Bu kısımda bir ya da daha çok bağımlı değişken olduğu durumlarda bağımlı değişkenlerin kestirimi ve regresyon katsayı tahmininin nasıl yapılacağı anlatılmıştır.

#### 2.2.3.1.4. Kısmi En Küçük Kareler Regresyon Modeli

KEKKR, çoklu doğrusal regresyon ve KEKK yöntemlerinin ikisini de içeren regresyon yöntemidir. Bu yöntemde boyut indirgenmesi ile elde edilen gizli değişkenler daha sonra açıklayıcı değişkenler olarak regresyon analizinde kullanılırlar. Yani gizli değişkenlerin elde edilmesi ve regresyon işlemi bütünlük vaziyettedir.

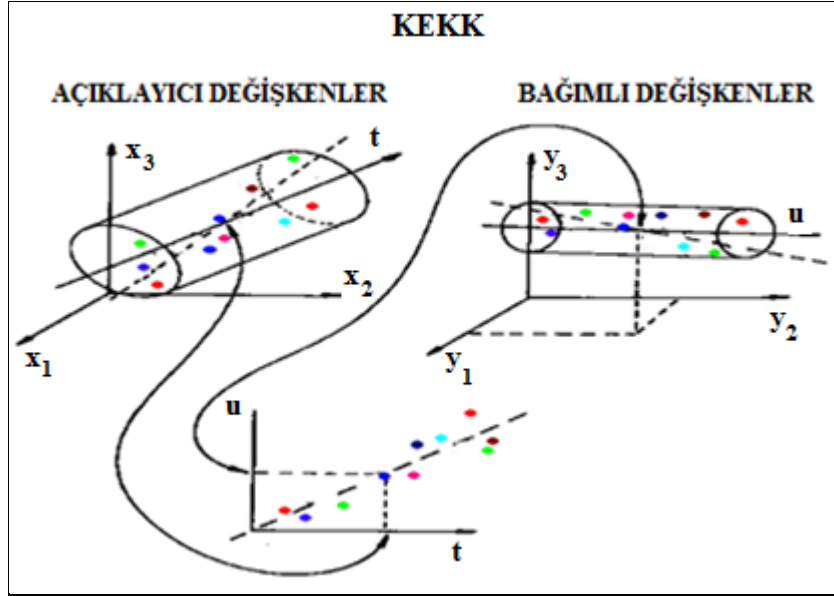
KEKKR yöntemi açıklayıcı ve bağımlı değişkenler arasındaki kovaryansı maksimum yapacak şekilde seçerek optimum sayıda gizli değişken sayısını elde eder. Bu yöntemde bağımlı ve açıklayıcı değişkenler için ayrı olarak modelleme işlemi yapılmaktadır.

$$X = \sum_{a=1}^A t_a p'_a + E = TP' + E \quad , \quad (a = 1, \dots, A) \quad (77)$$

$$Y = \sum_{a=1}^A u_a q'_a + G = UQ' + G \quad (78)$$

Yukarıdaki eşitliklerde  $X$ ,  $(nxp)$  boyutlu açıklayıcı değişkenler matrisi ve  $Y$ ,  $(nxm)$  boyutlu bağımlı değişkenler matrisi için oluşturulan modellerdir. Bu eşitliklerdeki  $T$  ve  $U$   $(nxA)$  boyutlu olup sırasıyla  $X$  ve  $Y$ 'nin gizli değişken (skor) matrisleridirler.  $t_a$  ve  $u_a$ 'lar ise  $(nx1)$  boyutlu olup sırasıyla  $X$  ve  $Y$  skor matrislerinin kolon vektörlerini meydana getiren gizli değişkenlerdir. Gizli değişkenler, bağımlı değişkenlerin tahmin edilmesinde kullanılmaktadırlar. Eşitliklerdeki  $A$ ,  $X$  skorlarını göstermek için kullanılmakta olup bu skorlar az sayıda ve birbirlerine diktirler.  $P'$ ,  $X$  yük matrisi olup  $(Axp)$  boyutludur ve  $(1xp)$  boyutlu  $p'_a$  ise  $X$  yüklerinin satır vektörüdür. Aynı şekilde  $Q'$ ,  $(Axm)$  boyutlu  $Y$  yük matrisi olup  $(1xm)$  boyutlu  $q'_a$  ise  $Y$  yüklerinin satır vektörüdür. Ayrıca  $(nxp)$  boyutlu  $E$ ,  $X$  bloğunun artık matrislerini,  $(nxm)$  boyutlu  $G$  ise  $Y$  bloğunun artık matrislerini oluşturmaktadırlar. Şekil 15 ile açıklayıcı ve bağımlı değişkenler

üzreinden projeksiyon yoluyla elde edilen gizli değişkenlerin temsili olarak ifadesi gösterilmiştir.



Şekil 15. KEKK'de gizli değişken yapıları üzerine projeksiyonun temsili olarak gösterilmesi [50].

Bu eşitliklerin aşama aşama oluşumu ve kullanılan parametrelerin anlamları ayrı ayrı incelenecek olursa;  $t_i$ 'ler  $x_j$  değişkenleri ve ağırlık olarak isimlendirilen  $w_{ja}^*$ 'ların doğrusal birleşimi olup (79) eşitliğindeki gibi kestirilirler. Eşitlikteki  $n$ , gözlem sayısını,  $p$  ise açıklayıcı  $X$  değişkenlerini gösteren göstergelerdir.

$$t_{ia} = \sum_j w_{ja}^* x_{ij} \quad , \quad (T = XW^*) \quad (79)$$

$$x_{ij} = \sum_a t_{ia} p'_{aj} + e_{ij} \quad , \quad (X = TP' + E) \quad (i = 1, \dots, n), (j = 1, \dots, p) \quad (80)$$

$$y_{il} = \sum_a u_{ia} q'_{al} + g_{il} \quad , \quad (Y = UQ' + G) \quad (l = 1, 2, 3, \dots, m) \quad (81)$$

Eşitliklerdeki  $p$  yük matrisi,  $e$  artık matrisidir.  $u_a$  ve  $q_{al}$ 'lar ise sırasıyla  $Y$ 'nin skorları ve ağırlıkları,  $m$ , bağımlı değişkenler için gösterge,  $g_{il}$  artıklardır.

$$y_{il} = \sum_a q'_{ia} t_{ia} + f_{il} \quad , \quad (Y = Q'T + F) \quad (82)$$

$$y_{il} = \sum_a q'_{la} \sum_j w_{ja}^* x_{ij} + f_{il} = \sum_a b_{lj} x_{ij} + f_{il} \quad (83)$$

$$b_{lj} = \sum_a q'_{la} w_{ja}^* \quad , \quad (B = W^*Q') \quad (84)$$

Açıklayıcı değişkenlerin skorları olan  $t$ 'ler (82)'deki gibi aynı zamanda bağımlı değişkenlerin tahmin edilmesinde de kullanılırlar. Bu eşitlikteki  $f$  artıkları ise modellenen bağımlı değişkenler ile gözlenen değerler arasındaki farktan oluşan artıklardır. Eşitlik (84) ise KEKKR modelindeki regresyon katsayılarını ifade etmektedir. Eşitliklerin yan taraflarında matris şeklindeki gösterimleri de mevcuttur.

KEKKR modeli alternatif olarak  $w_a$  cinsinden de ifade edilebilmektedir. Bu işlemde  $X$  matrisi yerine  $E_{a-1}$  artık matrisi kullanılır. Bu durumda skor vektörü (85)'deki gibi ifade edilir. (86) ise artıkların güncellenmesini belirtmektedir.  $Y$  matrisi için de (86) eşitliğine benzer şekilde  $t_a q'_a$  çıkartılarak indirgenme gerçekleştirilebilir. Bu yöntemde ilk KEKK bileşeni  $X$  matrisinden elde edilmektedir.  $w$  ağırlıkları ise doğrudan  $X$  ile bağımlı  $w^*$  yeni ağırlıklarına dönüştürülebilmektedir. Bu iki ağırlık arasındaki bağlantı  $W^* = W(P'W)^{-1}$  şeklindedir.  $W$ ,  $w_a$  kolonlarından meydana gelen  $X$  ağırlık matrisidir [51-52].

$$t_{ia} = \sum_j w_{ja} e_{ij,a-1} \quad (t_a = E_{a-1}w_a) \quad (85)$$

$$e_{ij,a-1} = e_{ij,a-2} - t_{i,a-1}p_{a-1,j} \quad (E_{a-1} = E_{a-2} - t_{a-1}p'_{a-1}) \quad (86)$$

$$e_{ij,0} = x_{ij} \quad (E_0 = X) \quad (87)$$

Yukarıda bahsi geçen  $w_a$  vektörleri birimdik,  $t_a$  vektörleri ise birbirlerine diktir.  $p_a$  ile  $u_a$ 'lar ise birbirlerine dik değıllerdir. Ayrıca  $u$ 'lar  $t$ 'lere ve  $p$ 'ler  $w$ 'lere diktir [51-52].

KEKK modeli dışsal ve içsel ilişkilerden meydana gelmektedir. İçsel ilişkiler  $X$  ve  $Y$  bloklarını birleştiren, dışsal ilişkiler ise ayrı ayrı her iki blok için olan ilişki olarak düşünülebilir.  $X$  ve  $Y$  blokları için sırasıyla dışsal ilişkiler (80) ve (81) eşitlikleriyle ifade edilir. Bu dışsal ilişkiler kullanılarak  $Y$  bloğunu olabildiği kadar açıklayabilmek ve  $\|G\|$ 'i olabildiğince küçük yapmak amaçlanır. İç ilişki ise her bir bileşen için  $u$  skorunun  $t$  skoruna karşı grafiğine bakılarak elde edilir. Bu ilişkinin doğrusal modeli  $\hat{u}_a = b_a t_a$  şeklindedir. Bu eşitlikteki  $b_a$  bir KEKK bileşeni için regresyon katsayısıdır ve  $b_a =$

$u'_a t_a / t'_a t_a$  'dır. (81) eşitliğindeki gibi bir model  $X$  ve  $Y$  matrisleri için temel bileşenler ayrı ayrı hesaplandığından ve bundan ötürü birbirleri ile zayıf bir ilişkiye sahip olduklarından elde edilebilecek en iyi model değildir. Bu sebeple daha karma bir ilişki olan  $Y = TBQ' + F$  modeli kullanılır. Bu modeldeki amaç da daha önce bahsedilen dış ilişki modeline benzer şekilde  $\|F\|$  'i en küçük yapmaktır. KEKKR yönteminde  $X$  matrisi ayrıştırılarak  $Y$  matrisi elde edilir. Bu işlem sırasında  $X$  skorları  $t_a = E_{a-1} w_a$  olarak tahmin edilir.  $X$  ve  $Y$  blokları için sırasıyla artık matrisleri ise  $E_a = E_{a-1} - t_a p'_a$  ve  $F_a = F_{a-1} - b_a t_a q'_a$  şeklindedir [34].

Herhangi bir modellemede kullanılacak olan bileşen sayısı KEKK modeli için önemli bir özelliktir. Çoklu bağlantı probleminden ötürü genellikle önemsiz bileşenler modelden çıkartılmaktadır. Fakat bu işlem gerçekleştirilirken hangi aşamada durulacağına karar vermek önemli bir noktadır. Bunun için bir takım yöntemler mevcuttur. Bu yöntemlerden birisi de  $F_a = F_{a-1} - b_a t_a q'_a$  eşitliğinden yararlanılarak elde edilir.  $F_a$  artıklarının normu küçük bir değere sahip olmalıdır. Bu norm ile bileşen sayısı arasındaki ilişkiyi gösteren bir grafik çizildiğinde  $\|F_a\|$  'nın önceden belirlenen bir değer altına düşmesi halinde grafikte karşılık gelen bileşen sayısı aranan bileşen sayısı olarak değerlendirilebilir. Diğer bir yöntem ise  $\|F_a\|$  'nın o andaki ve bir önceki değerine bakılarak aradaki farkın daha önceden belirlenmiş olan bir ölçüm hata değerinin altına düşmesi durumunda durma ile sağlanır. Bahsedilen bu yöntemlerin dışında birçok yöntem bileşen sayısını belirlemek için kullanılabilir [34,51,52].

#### 2.2.3.1.4.1. KEKKR Model Geçerliliğinin İncelenmesi

KEKKR yönteminde seçilen model yardımıyla elde edilen kestirim değerinin güvenilirliğinin test edilmesi için çeşitli ölçütler kullanılmaktadır. Bir modelin geçerliliğinin en iyi göstergesi elde edilen yeni  $X$  değerleri ile  $Y$  değerlerini gerçeğe en yakın şekilde kestirmektir. Bu çalışmada model geçerliliğinin incelenmesi için hata kareler ortalamasının karekökü (RMSE - root mean square error) kullanıldı. RMSE yöntemi modelin kullanılan veri kümesine uyumunu göstermektedir. RMSE eşitlik (88) ile hesaplanmaktadır.

$$RMSE(\hat{y}) = \sqrt{MSE(\hat{y})} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (88)$$

Eşitlikteki  $n$  gözlem sayısını,  $\hat{y}_i$ 'ler bütün gözlemler model bilgisinde bulunduğu kestirilen değişken değerlerini ifade etmektedir. RMSE değeri küçük olan modeller veriye daha iyi uyum sağlayan modellerdir.

### 2.2.3.1.5. Kısmi En Küçük Kareler Regresyonunda Algoritmalar

KEKKR yönteminde aralarında çoklu bağlantı sorunu olan açıklayıcı değişkenler çeşitli algoritmalar aracılığı ile açıklayıcı ve bağımlı değişkenlerdeki değişimi açıklayacak, çoklu bağlantı sorunu ortadan kalkmış açıklayıcı değişken sayısından daha az sayıdaki bileşene indirgenmektedir. KEKKR yönteminde kullanılan bu algoritmaların başlıcaları NIPALS ve SIMPLS algoritmalarıdır. Yapılan bu çalışmada kullanılacak olan algoritma ise SIMPLS algoritmasıdır. KEKKR yönteminde ilk kullanılan algoritma NIPALS algoritması olup SIMPLS'de bu algoritmanın geliştirilmesi ile elde edilmiştir. Bu yüzden NIPALS, KEKKR yönteminin temelini oluşturmaktadır.

#### 2.2.3.1.5.1. NIPALS Algoritması

Klasik algoritma olarak da bilinen NIPALS algoritması KEKK'in temelini oluşturmaktadır. Bu algortmada amaç kovaryans matrisini maksimize eden bileşenleri elde etmektir. Yinelemeli bir algoritma olan NIPALS'in işleyişi basamak basamak olup, her bir basamakta bir bileşen ve bu bileşene ait yük ve ağırlık değerleri bulunmaktadır. Algoritmanın sonlandırılması ise istenilen bileşen sayısına ulaşıncaya ya da  $X$  matrisi sıfır matrisine dönüştüğünde gerçekleşir. Algoritma çok sayıda ya da tek sayıda bağımlı değişkenin olduğu durumlarda kullanılabilir. Tek sayıda bağımlı değişkenin olduğu durumlarda farklı olarak algoritma yinelemeli olmaz. Algoritmanın basamak basamak ilerleyişi aşağıda numaralandırılarak gösterilmektedir. Bu adımlarda açıklayıcı değişken matrisi ( $n \times k$ ) boyutlu  $X$ , bağımsız değişken matrisi ise ( $n \times p$ ) boyutlu  $Y$ 'dir. Ayrıca  $k$ , açıklayıcı değişkenlerin sayısını,  $p$  bağımlı değişkenlerin sayısını,  $a$  ise bileşen sayısını belirtmektedir

1. Başlangıç basamağında  $X_1 = X$  ve  $Y_1 = Y$  olarak belirlenmektedir. Çok sayıda bağımlı değişkenin bulunduğu durumlarda  $u_a$  vektörü olarak  $Y$  matrisinin en



yüksek varyansa sahip olan sütunu veya ilk sütunu, bağımlı değişken sayısının tek olduğu durumlarda ise değişken sütununun kendisi seçilir.

2.  $X$  matrisinin  $u_a$  vektörü üzerine regresyonu ile  $X$  ve  $u_a$  arasındaki kovaryansı maksimize eden  $w$  ağırlık vektörü elde edilir. Yani  $w_a = X'_a u / (u'_a u_a)$  hesaplanır.
3.  $w_a$ 'nın ölçeklendirilmesi için  $w_a / \|w_a\|$  işlemi yapılarak boyu 1 olacak şekilde ölçekleme yapılır.
4.  $X$  skoru olan  $t_a$ ,  $w_a$  ağırlık vektörü ile  $t_a = X_a w_a$  olacak şekilde  $X$ 'in doğrusal kombinasyonu olarak elde edilir.
5.  $Y$ 'nin  $t_a$  üzerine regresyonu ile  $t_a$ 'nın  $Y$ 'i modellemedeki katkısı  $c_a = Y'_a t_a / (t'_a t_a)$  ile bulunur.
6.  $c_a / \|c_a\|$  ile  $c_a$ 'nın boyu 1 olacak şekilde ölçeklendirilir.
7.  $Y$ 'nin  $c_a$  ağırlık vektörü ile doğrusal kombinasyonundan  $Y$  skorlarının güncellenmiş durumu olan  $u_{a(yeni)} = Y_a c_a / (c'_a c_a)$  hesaplanarak  $Y$  için ilgili bileşen bulunur.
8. İkinci ve yedinci basamakta hesaplanan  $u_a$  değerleri arasında bir yakınsama olup olmadığı incelenir. Bunun için iki vektör arasındaki farkın norm değerinin sifira çok yakın bir sayıya eşit olup olmadığı incelenir. Eğer yakınsama sağlanamamışsa yedinci basamaktaki  $u_{a(yeni)}$  değeri ikinci basamakta yerine yerleştirilerek algoritmaya devam edilir. Yakınsamanın sağlandığı durumda ise dokuzuncu basamaktan tekrar algoritmaya devam edilir.
9.  $X$  matrisinin ilgili bileşeninin  $t_a$  skoru üzerine regresyonu ile  $p_a = X'_a t_a / (t'_a t_a)$  yük vektörü bulunur. Bu vektör bileşenin açıklayıcı değişken üzerindeki etkisini ifade eder.
10.  $Y$  matrisinin ilgili bileşeni  $u_a$  üzerine regresyonu ile  $q_a = Y'_a u_a / (u'_a u_a)$  yük vektörü bulunur. Bu vektör de bileşenin bağımlı değişken üzerindeki etkisini ifade eder.
11. Bu algoritmada açıklayıcı ve bağımlı değişkenler olan  $X$  ve  $Y$  için bileşenler ayrı ayrı hesaplandığı için bu bileşenler arasındaki ilişki güçsüz olmaktadır. Bu ilişkiyi daha kuvvetli hale getirmek için bileşenlerin her biri için  $u_a$ 'nın  $t_a$  üzerine regresyonundan  $b_a = u'_a t_a / (t'_a t_a)$  hesaplanır. Böylece içsel bir ilişki tanımlanmış olur.

12. Açıklayıcı ve bağımlı değişkenlerin modellenmesi için bulunan tüm yük ve bileşenler kullanılarak  $X = TP'$  ve  $Y = BTC'$  modelleri elde edilir. Sonrasında ise artık matrisler olan  $X_{a+1} \rightarrow X_a - t_a p'_a$  ve  $Y_{a+1} \rightarrow Y_a - b t_a c'_a$  hesaplanarak sonraki basamak bileşenini elde etmede kullanılır. Açıklayıcı değişkenler matrisinde bağımlı değişkenler hakkında daha fazla önemli bilgi kalmadığı anlaşılıncaya kadar algoritmaya ilk basamağa geri dönülerek devam edilir. Bu algoritma gerekli olan en az sayıdaki bileşen sayısını verir.

### 2.2.3.1.5.2. SIMPLS Algoritması

SIMPLS algoritmasındaki amaç NIPALS'den farklı olarak skorları, açıklayıcı değişken matrisinin indirgenmiş biçimi yerine, merkezleştirilmiş açıklayıcı değişken matrisinin doğrusal birleşimleri olarak ifade etmektir. Bu algorithmada bileşenler  $X$  ve  $Y$  arasındaki kovaryans matrisini maksimize edecek şekilde ve belirlenen normalleştirme ve diklik şartlarını yerine getirerek belirlenmektedir. SIMPLS ve NIPALS algoritması ile oluşturulan modeller arasındaki fark çok küçük olmakla birlikte oluşturulan her iki model birbirine benzemektedir. SIMPLS'de önceki bileşenlerden oluşan yük vektörlerine yansıtılan merkezleştirilmiş  $X'Y$  matrisi üzerinde özvektör analizi yapılır [53]. NIPALS algoritmasında olduğu gibi açıklayıcı ve/veya bağımlı değişken matrislerini indirgeme gerekmemektedir. Bu durumun faydası ise daha az hafıza ve hesaplama gerektirmesidir. Ayrıca merkezleştirilmiş açıklayıcı değişken matrisinin doğrusal birleşimleri olarak bütün bileşenleri analiz etmek daha kolaydır. Bu algorithmada öncelikli olarak hedef belirlenip bu hedefe en uygun ölçüt bulunur. Sonrasında bulunan ölçüt denenerek uygun hale dönüştürülür. Tüm bu işlemler sonrasında ise algoritma kurulur [54].

SIMPLS algoritması iteratif bir algoritma olup algoritmanın amacı KEKKR modeli olan  $\hat{Y} = X\hat{\beta}_{PLS}$ 'nin kestirimini yapmaktır. Bu modelin oluşturulması için aşağıdaki eşitlikte ifade edilen  $T$  ve  $U$  gizli değişkenlerine ihtiyaç duyulmaktadır [55].

$$T = XR, U = YC \quad (89)$$

Eşitlikte de görüldüğü gibi  $T$ , açıklayıcı değişken matrisi  $X$ 'in ağırlığı olan  $R$  ile doğrusal kombinasyonundan,  $U$  ise bağımlı değişken matrisi  $Y$ 'nin ağırlığı olan  $C$  ile

doğrusal kombinasyonundan meydana gelmektedir.  $T = (t_1, \dots, t_A)$  gizli değişken matrisi, ortogonal gizli değişkenler olan  $t_a = Xr_a$ 'lardan oluşmaktadır ( $t_i \perp t_j, i \neq j$ ) ve bu gizli değişkenler  $u_a = Yc_a, (a = 1, 2, \dots, A), (u_j \perp t_j, i < j \leq a)$  şeklindeki  $Y$  bileşenleri ile kovaryanslarını maksimize edecek şekilde  $S = X'Y$  kovaryans matrisinden belirlenirler. Bu maksimizasyon işlemi SIMPLS algoritmasında sırasıyla  $X$  ve  $Y$ 'nin ağırlık vektörleri olan  $r_a$  ve  $c_a$ 'a ( $a=1, \dots, A$ ) bağlı olarak gerçekleştirilir [56].

Orjinal  $X$  ve  $Y$  matrisleri ile başlayan algoritmayı basamaklayarak ifade edecek olursak iterasyon basamakları aşağıdaki şekilde gerçekleşir.

1.  $S_1 = X'Y$  kovaryans matrisi algoritmanın başlangıcında hesaplanır.
2.  $r_a, S_a S_a'$  vektörünün baskın özvektörü olarak elde edilir.
3.  $t_a = Xr_a$  hesaplanır.
4.  $t_a$  normalize edilir  $t_{a-norm} = \frac{t_a}{\sqrt{t_a' t_a}}$  ve  $X$  değişkenlerinin ağırlıkları  $t_{a-norm}$ 'a adapte edilir  $r_{a-adapt.} = \frac{r_a}{\sqrt{t_a' t_a}}$ .
5.  $p_a = X' t_{a-norm}, c_a = Y' t_{a-norm}$  ve  $u_a = Yc_a$  hesaplanır.
6. Kovaryans matrisi  $S$ 'nin güncellenmesi için  $S_{a+1} = S_a - v_a(v_a' S_a)$  eşitliği uygulanır. Bu eşitlikteki  $v_a = p_a - V_{a-1}(V_{a-1}' p_a)$  ( $a = 2, 3, \dots, A$ ) şeklinde tanımlanır. İlk iterasyonda  $v_1 = p_1$ 'dir.  $V_a \equiv (v_1, \dots, v_a)$ ,  $X$  yükleri olan  $P_a$ 'nın birim dik ortanormalleridir.  $P_a = (p_1, \dots, p_a)$ .
7.  $u_a, t_b$  ( $b = 1, 2, 3, \dots, a - 1$ ) ile ortogonal olup  $u_a = u_a - T_{a-1}(T_{a-1}' u_a)$  ve  $T_{a-1} \equiv (t_1, \dots, t_{a-1})$ 'dir. Bu basamak sadece  $a > 1$  için geçerlidir.
8. Sonraki  $t_{a+1}$  gizli değişkenleri,  $S_{a+1}$  kovaryans matrisi kullanılarak ikinci basamaktan itibaren elde edilmeye devam eder. Bu iterasyon bütün gizli değişkenler elde edilinceye kadar sürer. İterasyon sayısı gizli değişken sayısı olan  $A$ 'a eşittir. Sonrasında ise regresyon katsayıları hesaplanır. NIPALS algoritmasına benzer olarak SIMPLS algoritmasında da  $T = XR$  ve  $B = WC'$  eşitlikleri kullanılır.

### 2.2.3.2. k-En Yakın Komşuluk (k-NN) Yöntemi

k en yakın komşuluk yöntemi sınıflandırma problemlerinde kullanılan denetimli öğrenme algoritmalarından birisidir. Bu yöntemde kullanılan algoritma ile test örneğinin

eđitim kümesinde bulunan tüm örneklere olan yakınlığı seçilen bir uzaklık ölçütüne göre hesaplanır. Yapılan bu çalışmada uzaklık ölçütü olarak öklid uzaklığı tanımlanmıştır. Hesaplanan uzaklık değerleri küçükten büyüğe doğru sıralanır ve en yakın olan  $k$  adet örnek bulunup bu örnekler en çok hangi sınıfa ait ise test örneđi de o sınıfa dahil edilir. İki nokta arasındaki öklid uzaklığının hesaplanması için eşitlik (90) ile ifade edilen formül kullanılmaktadır. Formüldeki  $R$  özellik sayısı,  $p$  ve  $q$  ise uzayındaki noktaları ifade etmektedir.

$$d_E(p, q) = \sqrt{\sum_{i=1}^R (p_i - q_i)^2} \quad (90)$$

$k$ -NN algoritmasında kullanılan adımları sıralayacak olursak;

1.  $k$  değeri belirlenir
2. Test örneđinin tüm eğitim örnekleri ile olan uzaklığının hesaplanması
3. Minimum mesafeye göre sıralama işleminin gerçekleştirilmesi ve en yakın  $k$  komşuluğun belirlenmesi
4. Yakın komşulukların sınıf değerlerinin bulunması
5. Değeri baskın olan sınıfın örüntü sınıfı olarak belirlenmesi

$k$  en yakın komşuluk algoritması uygulanabilirliği basit olan ve gürültüye karşı dirençli bir algoritma olmasına karşın her bir test örneđinin tüm eğitim örneklerine olan uzaklıklarını hesaplama gerekliliđi açısından maliyeti yüksek ve yavaş bir algoritmadır. Öznitelik vektörlerinin boyutlarının azaltılması ile yapılacak olan uzaklık hesabı azalacağından yöntemin hızı biraz daha arttırılabilir.

Bu çalışmada  $k$  en yakın komşuluk yönteminde yüksek doğrulukta sınıflandırma tahminini sağlayan en uygun  $k$  parametresinin belirlenmesi için çapraz doğrulama işlemi ile eğitim verileri alt eğitim ve onaylama verileri olmak üzere iki kümeye ayrılmaktadır. Bu ayrılma işlemi için çeşitli metotlar kullanılmaktadır. Çalışmada kullanılan metod KFOLD metodudur.

KFOLD yöntemine göre ilk önce bir fold değeri belirlenir ve eğitim verileri seçilen bu fold değeri kadar alt gruba ayrılır. Her bir alt grup sırasıyla onaylama grubu olarak atanırken geriye kalan onaylama grubu olmayan gruplar ise alt eğitim grubu olarak seçilirler. Böylece tüm eğitim alt grupları farklı basamaklarda onaylama ve alt eğitim

olacak şekilde atanmış olur. Çalışmada bu yöntem kullanılarak uygulanan algoritma ile en uygun k değeri belirlenmiştir.

### 2.2.3.3. Bayes Yöntemi

Bayes yöntemi Thomas Bayes tarafından geliştirilen ve koşullu olasılıkların hesaplanmasında kullanılan bir yöntemdir. İstatistiksel bir sınıflandırıcı olan bayes sınıflandırma ise bayes teoremine dayanan basit bir yöntem olup örüntülerin hangi sınıfa hangi olasılıkla ait olduklarını öngören bir sınıflandırma yöntemidir. Bayes yöntemi bütün sınıflandırma işlemleri için bütün veri kümesini işler. Kolay uygulanabilir ve birçok durumda iyi sonuçlar veren bir sınıflandırma yöntemidir. Değişkenler arası ilişki ise modellenemez.

Bayes teorisi sınıflandırma işlemine bir olasılık problemi olarak yaklaşmaktadır. Bayes teoreminin ispatı için koşullu olasılık tanımından başlanılır.

Bayes teoremi bir durum sırasında rastgele olarak ortaya çıkan A ve B olayları için koşullu ve marjinal olasılıklar arasındaki ilişkidir. B olayı bilindiğinde A olayının ve A olayı bilindiğinde B olayının olasılık ifadeleri sırasıyla (91) ve (92) eşitliklerinde ifade edilmiştir. Bu iki denklem düzenlendiğinde ise (93) eşitliği elde edilir.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (91)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (92)$$

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A) \quad (93)$$

Eşitlik (93)'ün her iki tarafı  $P(B)$  olasılığına bölüldüğünde (94) eşitliği ile ifade edilen bayes teoremi elde edilir. Teoremdeki her bir terimin ne anlama geldiği aşağıda açıklanmıştır.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} \quad (94)$$

$P(A|B)$ : B olayı gerçekleştiğinde A olayının meydana gelme olasılığı (B için A'nın koşullu olasılığı)

$P(B|A)$ : A olayı gerçekleştiğinde B olayının meydana gelme olasılığı (A için B'nin koşullu olasılığı)

$P(A)$ : A olayının önsel olasılığı (B olayı hakkında önceden herhangi bir bilgi içermediği için.)

$P(B)$ : B olayının önsel olasılığı (B'nin marjinal olasılığı)

Bayes sınıflandırma mekanizması ve bu mekanizmada kullanılacak olan büyüklükler aşağıdaki gibi tanımlanır.

$\Omega = \{w_1, w_2, \dots, w_s\}$  : s adet tanınması istenen örüntü sınıf kümesi (sonlu doğal durum kümesi)

$A = \{\alpha_1, \alpha_2, \dots, \alpha_a\}$  : a adet sonlu mümkün olabilen karar kümesi

$\underline{x} \in \mathfrak{R}^d$  : d bileşenli özellik/örüntü vektörü

$p(\underline{x}|w_j)$  :  $\underline{x}$  için şartlı olasılık yoğunluk fonksiyonu

$p(w_j|\underline{x})$  : Sonsal olasılık

$p(w_j)$  : Doğal  $w_j$  durumunda olan öncesel olasılık

Olasılıklar bayes kuralı ile birleştirildiğinde (95) ve (96) eşitlikleri elde edilir,

$$p(w_j|\underline{x}) = \frac{p(\underline{x}|w_j)p(w_j)}{p(\underline{x})} \quad (95)$$

$$p(\underline{x}) = \sum_{j=1}^S p(\underline{x}|w_j)p(w_j) \quad (96)$$

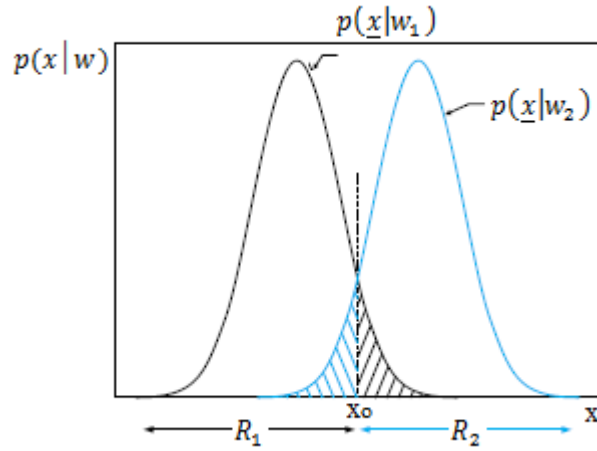
Bayes kuralına göre  $p(w_i|\underline{x}) > p(w_j|\underline{x})$  ise karar  $w_i$  olarak verilir ( $\forall j \neq i$ ). Bayes sınıflandırıcının birçok çeşitlemesi vardır fakat esas yapı aynı olup örüntü tanıma için yeterlidir [57].

Örüntü sınıflandırıcıları ifade etmek için en kullanışlı olan ve geniş kullanıma sahip olan yollardan birisi de diskriminant fonksiyonunun kullanımınıdır. Diskriminant fonksiyonunu tanımlayacak olursak;

Örneğin elimizde bir x verisinin var olduğunu ve bu verinin önceden tanımlanmış olan sınıflardan hangisine ait olduğunu bulmak istediğimizi düşünelim. x verildiği zaman onun  $w_i$  sınıfına dahil olma olasılığı aşağıdaki şekilde ifade edilir.

$$p(w_j|\underline{x}) = \frac{p(\underline{x}|w_j)p(w_j)}{p(\underline{x})} \quad (97)$$

Kaç adet sınıf mevcut ise bütün sınıflara göre  $p(w_j|\underline{x})$  ifadesi hesaplanır ve hangisinin olasılığı daha büyük ise  $x$ , o sınıfa dahil edilir. İki sınıf olduğunu ve veri kümelerinin dağılımlarının gaus şeklinde olduğunu düşünürsek her bir sınıfın olasılık yoğunluk fonksiyon grafiği Şekil 16'a benzer olacaktır.



Şekil 16. İki sınıfa ait olasılık yoğunluk fonksiyon grafiği

Şekildeki  $x_0$  ayrılma noktasıdır. Eğer  $w_1$ 'in olasılığı fazla ise bu nokta sağa, az ise bu nokta sola kayacaktır. Şekilde  $R_1$  ve  $R_2$  uzunlukları ile bu kayma izlenebilir. Eğer olasılıklar eşit ise  $x_0$ 'ın sağındaki noktalar ile solundaki noktaların olasılıkları eşit olacaktır. Diskriminant ayırma fonksiyonunu belirlemek için (97) eşitliğindeki ifadenin her iki tarafının  $\ln$ 'i alınır;

$$g_i(x) = \ln p(\underline{x}|w_i) + \ln P(w_i) \quad (98)$$

$$g_i(x) > g_j(x) \quad j \neq i \quad (99)$$

(99) eşitliğine göre hangi sınıfa ait diskriminant fonksiyonu büyük ise veri o sınıfa dahil edilecektir. Yukarıdaki açıklamalarda elimizde iki sınıf olduğunu düşündüğümüz için tekil diskriminant fonksiyonu  $g(x) \equiv g_1(x) - g_2(x)$  şeklinde olacaktır. Bu ifade  $g(x) >$

0 şeklinde elde edilirse  $x$  verisi  $w_1$  sınıfına,  $g(x) < 0$  şeklinde edilirse  $w_2$  sınıfına karar verilir. Eğer  $p(\underline{x}|w_i)$  yoğunluk fonksiyonu çok değişkenli normal dağılımlı ise (98) ifadesi aşağıdaki hale dönüşecektir.

$$g_i(x) = -\frac{1}{2}(x - \mu_i)' \Sigma_i^{-1}(x - \mu_i) - \frac{K}{2} \ln 2\pi - \frac{1}{2} \ln |\Sigma_i| + \ln P(w_i) \quad (100)$$

Kovaryans matrisi  $\Sigma_i$ 'nin her bir farklı kategori için farklı olduğu durumlarda diskriminant fonksiyonu doğal kuadratik olacaktır ve genel ifadesi (101) eşitliğindeki gibi olur.

$$g_i(x) = x'W_i x + w_i'x + w_{i0} \quad (101)$$

$$W_i = -\frac{1}{2} \Sigma_i^{-1} \quad (102)$$

$$w_i = \Sigma_i^{-1} \mu_i \quad (103)$$

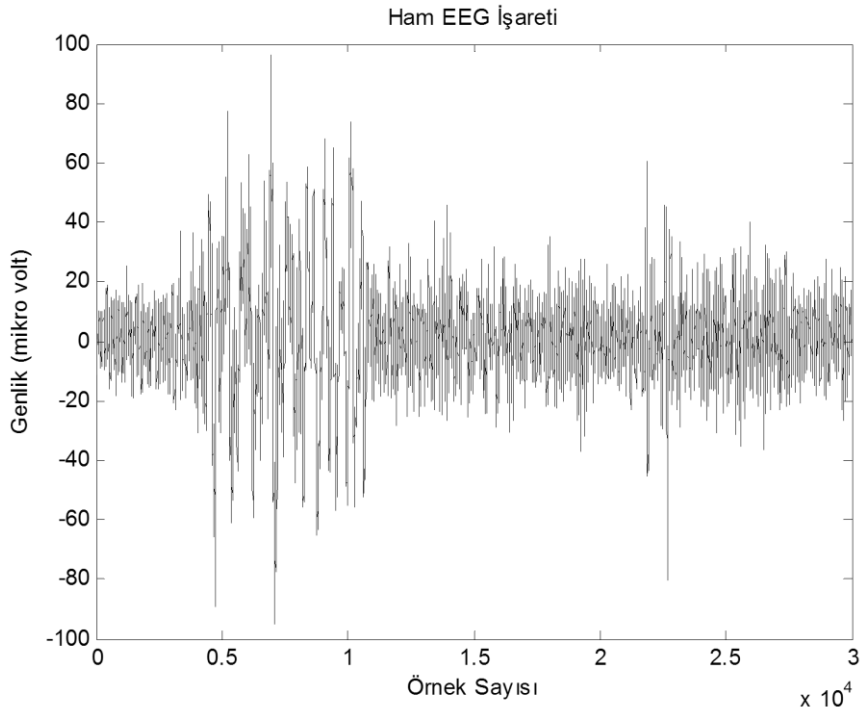
$$w_{i0} = -\frac{1}{2} \mu_i' \Sigma_i^{-1} \mu_i - \frac{1}{2} \ln |\Sigma_i| + \ln P(w_i) \quad (104)$$

Eşitliklerdeki  $\mu_i$  ortalama vektör,  $\Sigma_i$  ise kovaryans matrisidir [58].  $g_1$  ve  $g_2$  yüzey fonksiyonlar olduğundan ötürü  $g_i(x)$  diskriminant fonksiyonu bulunarak iki sınıf arasındaki sınırı elde etmiş oluruz. Bunların ortak bölgeleri ise ayırma fonksiyonudur.



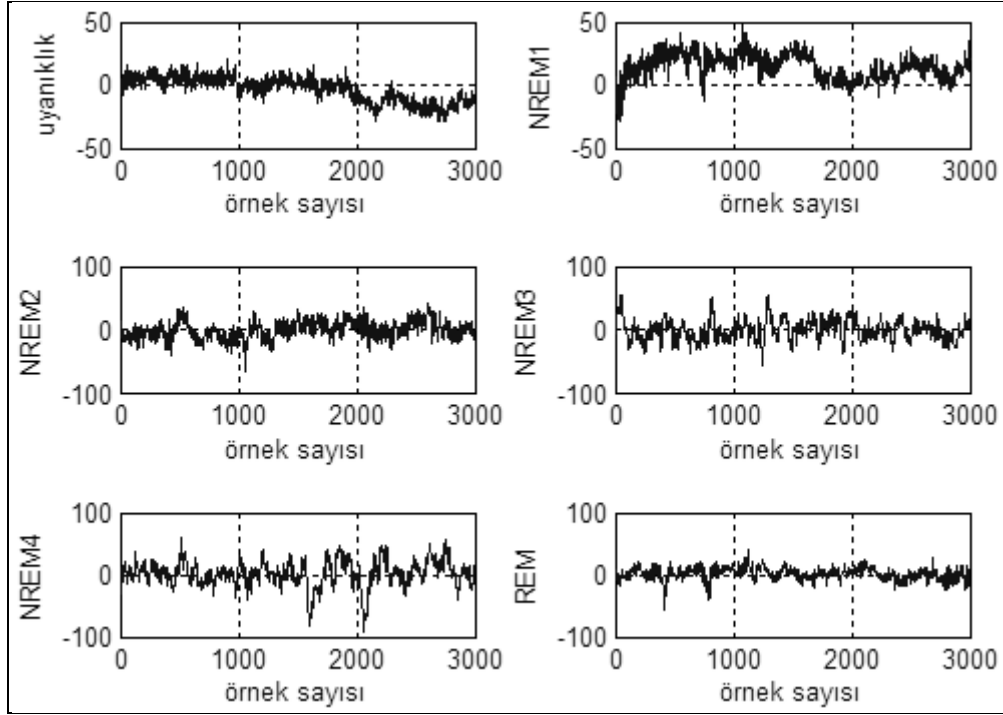
### 3. BULGULAR VE SONUÇLAR

Çalışmada kullanılan EEG kayıtları kafatasının 'PzOz' bölgesinden elde edilmiş olup bu kayıtlara ait özellikler ve kayıtlara uygulanan ön işlemler daha önce açıklanan materyal ve ön işleme bölümlerinde ayrıntısıyla verilmiştir. Tasarlanan algoritma için ilk olarak physionet veri tabanından elde edilen 'rec.'formatındaki EEG kayıt örnekleri 'mat.'formatına dönüştürülmüştür. Şekil 17 ile sağlıklı bir bireye ait, 'mat.'formatına dönüştürülmüş ham EEG kayıt örneği gösterilmektedir. Veri tabanından elde edilen EEG işaretleri birkaç saat süren çok uzun süreli kayıtlar olduklarından ötürü görsel açıdan daha anlaşılır ifade edilebilmeleri adına şekilde bu kayıtlara ait 5 dakikalık kayıt örneği kullanılmıştır. Veri tabanındaki kayıtlar 'mat.'formatına dönüştürüldükten sonra bütün EEG kayıtları şekil 18 ile gösterilen kişiye ait hipnogram grafiklerinden faydalanılarak uyanıklık ve uykunun beş ayrı evresine ayrılmıştır.



Şekil 17. Sağlıklı bir bireye ait beş dakikalık ham EEG kayıt örneği





Şekil 19. Sağlıklı bir bireyin uyanıklık ve uyku evrelerine ait epok örnekleri

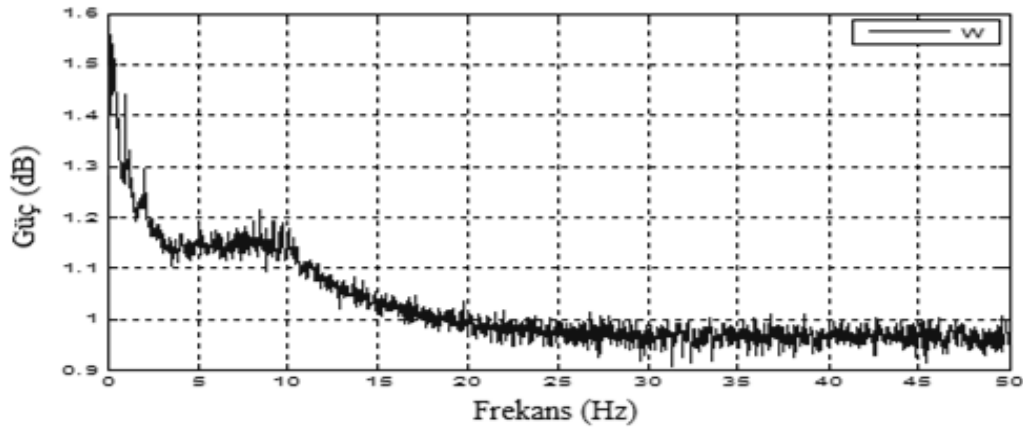
EEG işaretleri çok miktarda bilgi içeren işaretler olup bu nedenle bu işaretlerin analiz ve sınıflandırma işlemlerinin görsel olarak gerçekleştirilmesi zordur. Değişken yapıdaki EEG işaretleri zamanın ve frekansın bir fonksiyonu şeklindedirler. Bu sebeple EEG işaretlerinin zaman ve genlik düzlemlerinde analizleri gerçekleştirilebilmektedir. Frekans düzleminde gerçekleştirilen analiz spektrum analizi olarak adlandırılmaktadır. Rastgele işaretlerin enerjileri sonsuz olduğu için bu işaretlerin güç spektrumları incelenmektedir. İşaretlerin frekans bileşenlerinin güç yoğunluklarını belirlemek için çeşitli yöntemler kullanılabilir. Bu çalışmada hızlı fourier dönüşümünden yararlanılmıştır. Hızlı fourier dönüşümünün temeli de fourier dönüşümüne dayanır. Ayrık zamanlı bir işaretin güç spektrumu eşitlik (105)'deki gibi ifade edilmektedir.

$$\hat{P}(f_k) = \frac{f_s |X(k)|^2}{N} \quad (105)$$

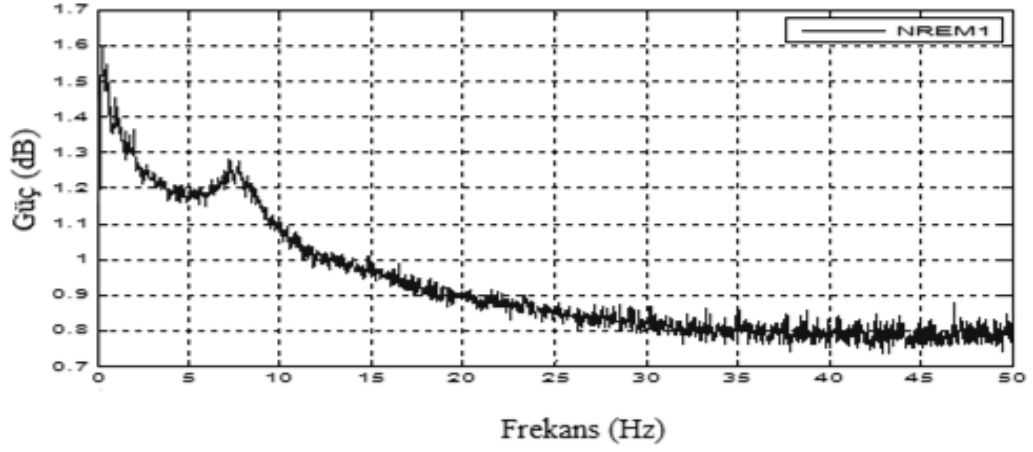
Bu denklemdeki  $f_s$  örnekleme frekansı,  $N$  örnek sayısı,  $X(k)$  ise ayrık zamanlı  $x(n)$  işaretinin ayrık fourier dönüşümüdür ve denklem (106) ile gösterilir.

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} \quad (k = 0, 1, \dots, N-1) \quad (106)$$

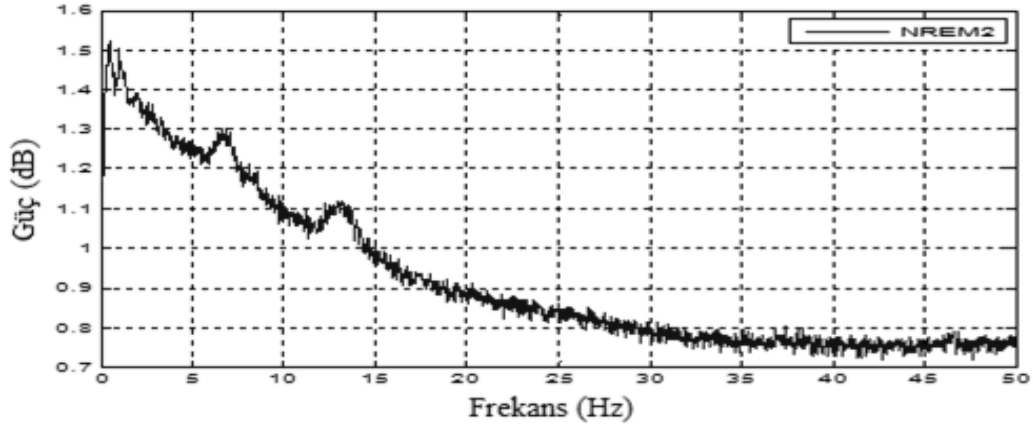
Yapılan bu çalışmanın öncesinde EEG verilerinin alt band frekans bileşenlerinin uyku ve uyanıklık durumu ile ilişkisini analiz etmek amacıyla güç spektrumları incelenmiştir. İnceleme sonucunda uyanıklık ile beş ayrı uyku evresinin (NREM1, NREM2, NREM3, NREM4, REM) kendi içlerinde farklı karakteristik özelliklere sahip oldukları, uyku ve uyanıklığın farklı evrelerinde farklı oranlarda baskın olarak beliren delta, teta, alfa ve beta frekans banlarındaki değişimin ne şekilde olduğu gözlemlenmiştir. Beta frekansındaki en yüksek aktivitenin uyanıklık evresinde olduğu, bunun yanında uykunun REM evresinde de bu aktivitede önceki evreye göre artış olduğu gözlemlenmiş, alfa aktivitesinden teta aktivitesine dönüşümün uyanıklık-uyku arası geçiş döneminde, teta ve delta aktivitelerindeki yoğun artışın ise derin uyku döneminde genel olarak daha baskın olduğu görülmüştür. Şekil 20-25 ile verilen grafiklerde veri tabanındaki bir kişiye ait EEG kayıtlarından kişinin sırasıyla bütün uyku evrelerinden elde edilen 30 saniyelik epokların ortalama güç-frekans spektrumları alt alta grafikler şeklinde verilmiştir. Ham EEG işaretleri yüksek oranda gürültü içerdiklerinden her uyku evresine ait 50 farklı epokun güç ortalamaları alınarak evrelerdeki frekans değişimlerinin daha iyi gözlemlenmesi sağlanmıştır. Bu grafikler incelendiğinde her evrede baskın olan frekans aralıkları yukarıda bahsedilen açıklamalara uygun olarak gözlemlenebilmektedir.



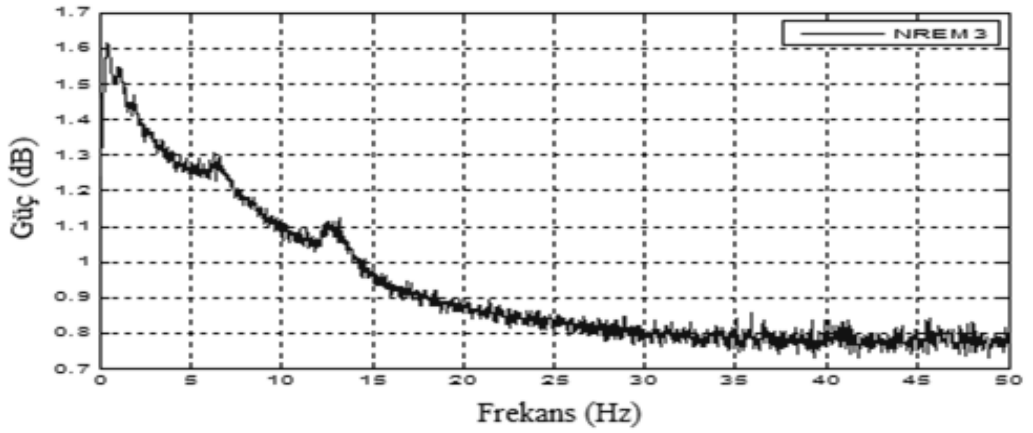
Şekil 20. W (uyanıklık) evresine ait ortalama güç-frekans grafiği



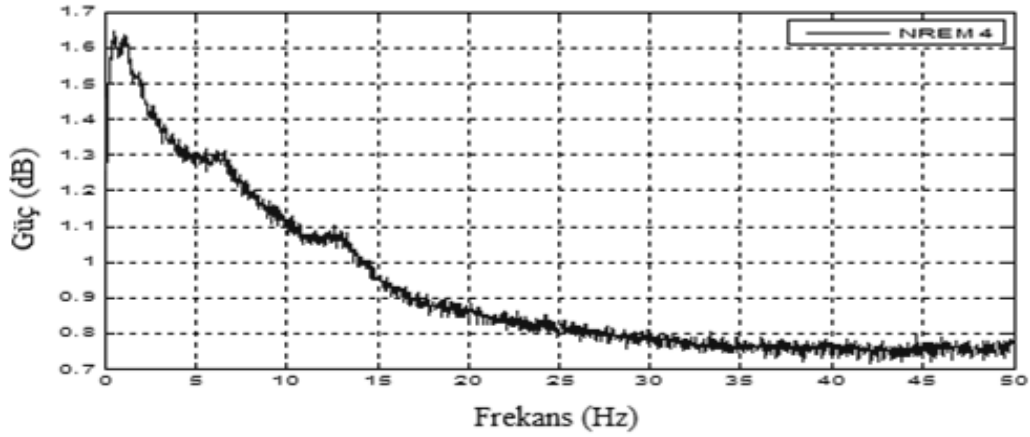
Şekil 21. NREM1 evresine ait ortalama güç-frekans grafiği



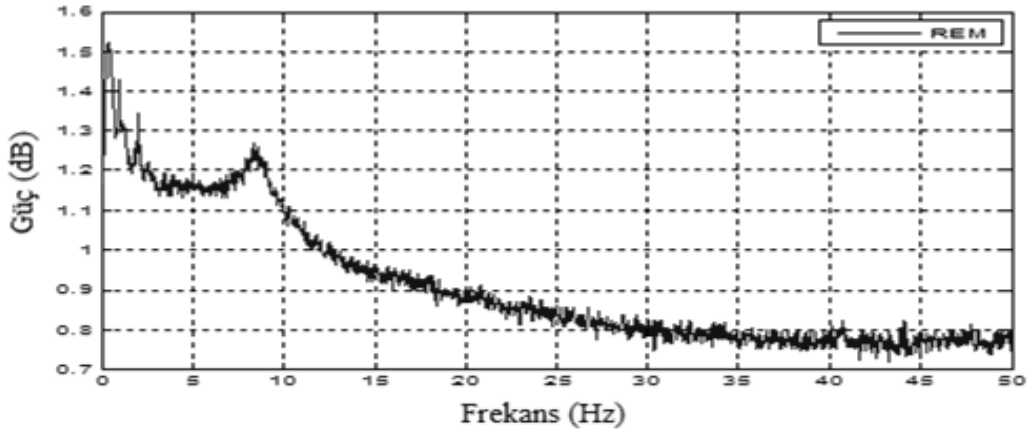
Şekil 22. NREM2 evresine ait ortalama güç-frekans grafiği



Şekil 23. NREM3 evresine ait ortalama güç-frekans grafiği



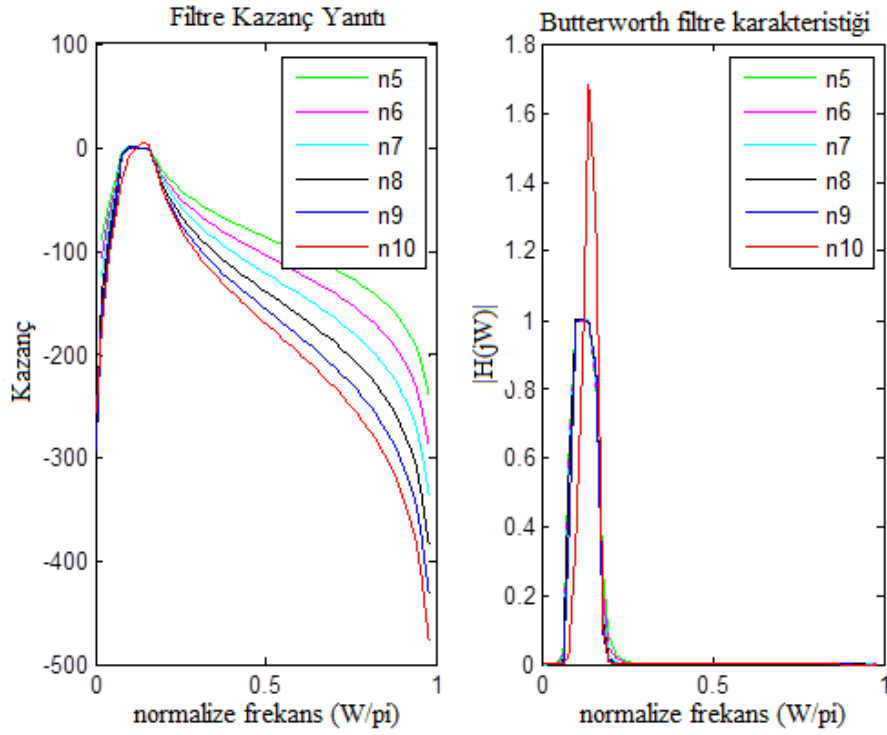
Şekil 24. NREM4 evresine ait ortalama güç-frekans grafiği



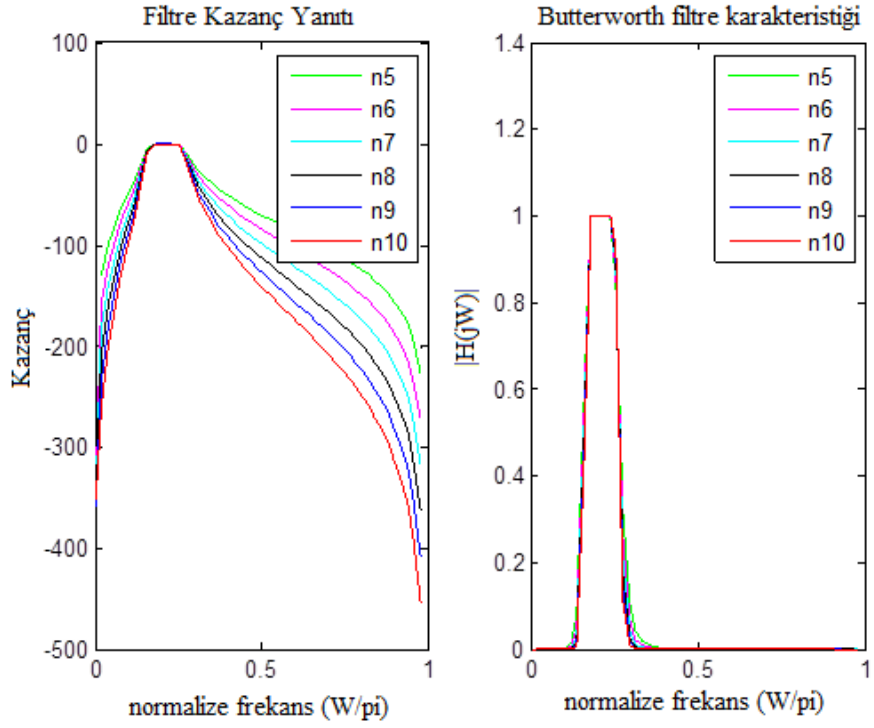
Şekil 25. REM evresine ait ortalama güç-frekans grafiği

Analiz işlemi sonrasında algoritma için gerekli verileri sağlamak amacıyla ham EEG kayıtlarının normalizasyonu gerçekleştirilerek EEG işaretleri, ortalaması sıfır, maksimum değerleri 1 olacak şekilde ölçeklendirilmişlerdir. Normalizasyon işlemi ardından elde edilen EEG epogu ile ham EEG epogunun karşılaştırılmasına ait grafik ön işleme bölümünde daha önceden sunulmuştur. Bu işlem sonrasında odaklanılması gereken frekans bandlarının elde edilebilmesi için incelenmek istenen her bir frekans aralığına ait band geçiren filtre tasarlanıp filtreleme işlemi gerçekleştirildi. Bu işlem için butterworth band geçiren filtre kullanıldı. Çalışmada hedeflenen şey uyku ve uyanıklığın sınıflandırılması olduğu için EEG kayıtlarında uyku başlangıcının yakalanması yani erken uyku döneminin belirlenmesi önem taşımaktadır. Daha önceki bölümlerde açıklandığı gibi uyanıklıktan uyku dönemine geçişin en büyük belirteçlerinden olan alfa ritminin yerini teta ritmine

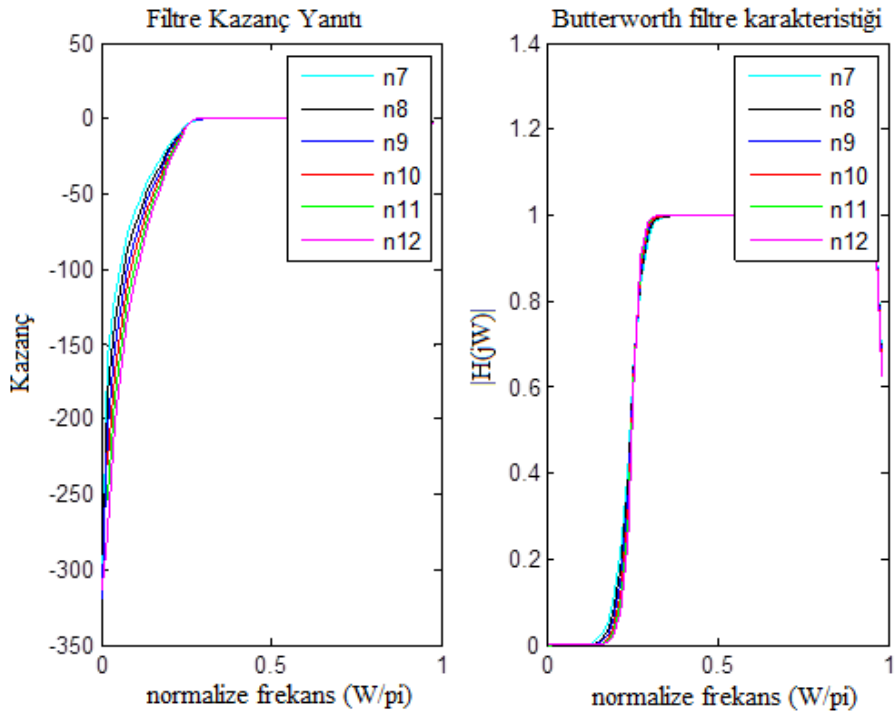
bırakması, gözlerin açık ve kapalı olma durumuna göre alfa ritminde değişmelerin meydana gelmesi, zihinsel işlev ve yoğun dikkat halinde beta ritminde artışların görülmesi ve bunun yanında delta dalgalarının ise derin uyku döneminde belirmesi gibi bilgilerden yola çıkılarak; çalışmada teta (4-8Hz), alfa (8-13Hz) ve beta (13-50 Hz) bandlarına ait band geçiren filtreler kullanılmıştır. Algoritmamızda oluşturulan filtrelerin her üç band aralığı ve farklı derece değerleri için kazanç yanıt ve genlik karakteristiklerine ait grafikler şekil 26-28 ile gösterilmiştir. MATLAB ortamında oluşturulan butterworth filtrelerin kazanç tepkeleri ve genlik karakteristikleri incelenerek en uygun filtre derecesi seçilip istenilen özelliklerde filtreler elde edildi.



Şekil 26. Teta band aralığı için farklı filtre derecelerine ait (n5, n6, n7, n8, n9, n10) kazanç yanıtı ve butterworth genlik karakteristikleri



Şekil 27. Alfa band aralığı için farklı filtre derecelerine ait (n5, n6, n7, n8, n9, n10) kazanç yanıtı ve butterworth genlik karakteristikleri



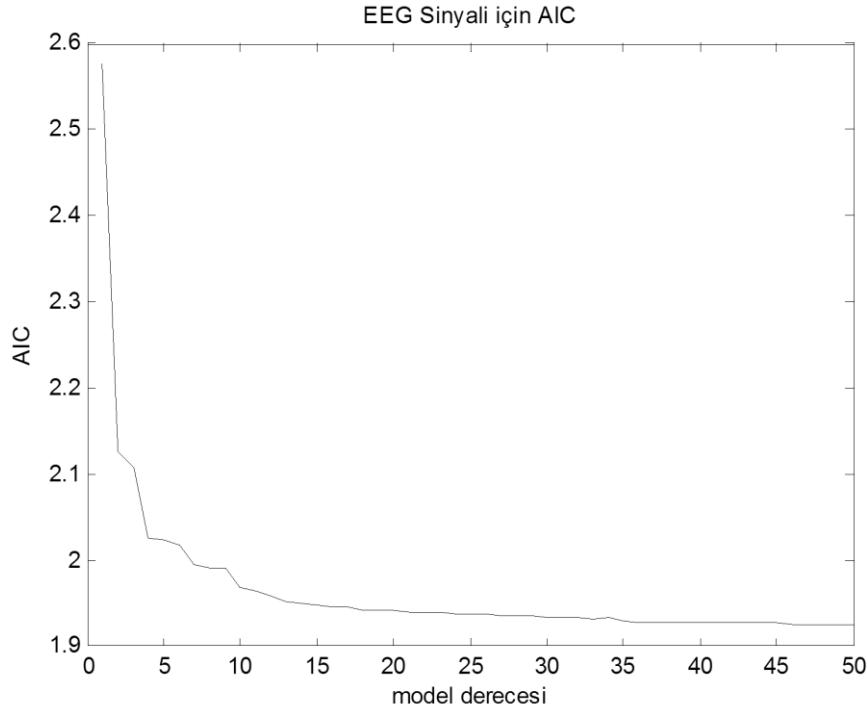
Şekil 28. Beta band aralığı için farklı filtre derecelerine ait (n7, n8, n9, n10) kazanç yanıtı ve butterworth genlik karakteristikleri



Oluşturulan bu üç filtre her kişi için tüm EEG verilerine ayrı ayrı uygulandı ve her kişi için üç filtre çıkışındaki EEG verileri kaydedildi.

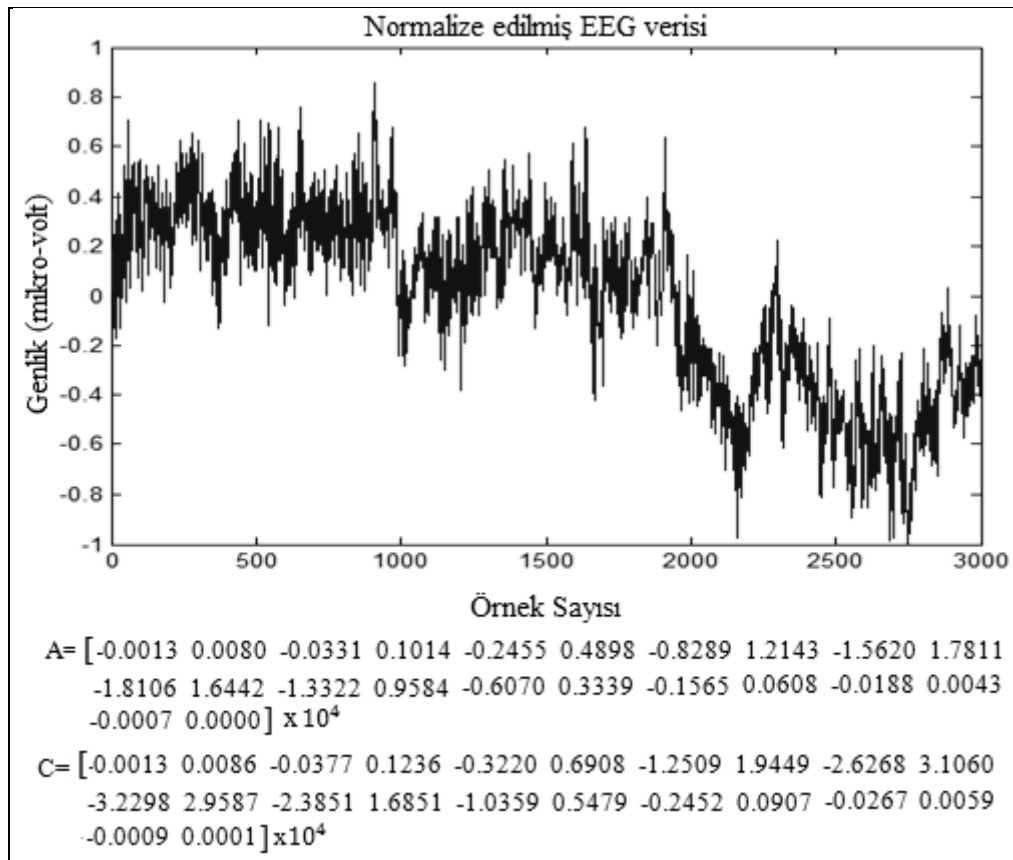
### 3.1. Öz Niteliklere Ait Bulgular

Filtreleme işlemi ardından her kişi için elde edilen teta (4-8Hz), alfa (8-13Hz) ve beta (13-50Hz) frekans bandlarına ait EEG verilerinin her birisi için AR yöntemi ile uyku ve uyanıklık dönemlerine ait öznelik vektörleri oluşturuldu. Algoritma sonucunda elde edilen AR katsayılarının adedi model derecesi ve kullanılacak olan algoritmanın çeşidi, sonraki aşama olan sınıflandırma sonucundaki başarı oranını etkileyen parametrelerdendir. Örneğin model derecesinin çok büyük seçilmesi EEG spektrumunda hatalı bölgelerin oluşmasına ve spektrumun bozulmasına, çok küçük seçilmesi ise spektrumdaki keskin bölgelerin net seçilmemesine sebep olabilmektedir. Bu durumda EEG spektrumunun AR katsayıları ile sağlıklı bir şekilde temsil edilmesini etkilemektedir. Bu çalışmada model derecesinin belirlenmesi için daha önceki bölümlerde bahsedilmiş olan AIC kriterinden yararlanılmıştır. AIC kriterinin EEG veri parçasına uygulanması sonucunda elde edilen grafik şekil 29 ile gösterilmiştir.



Şekil 29. EEG sinyali için AIC

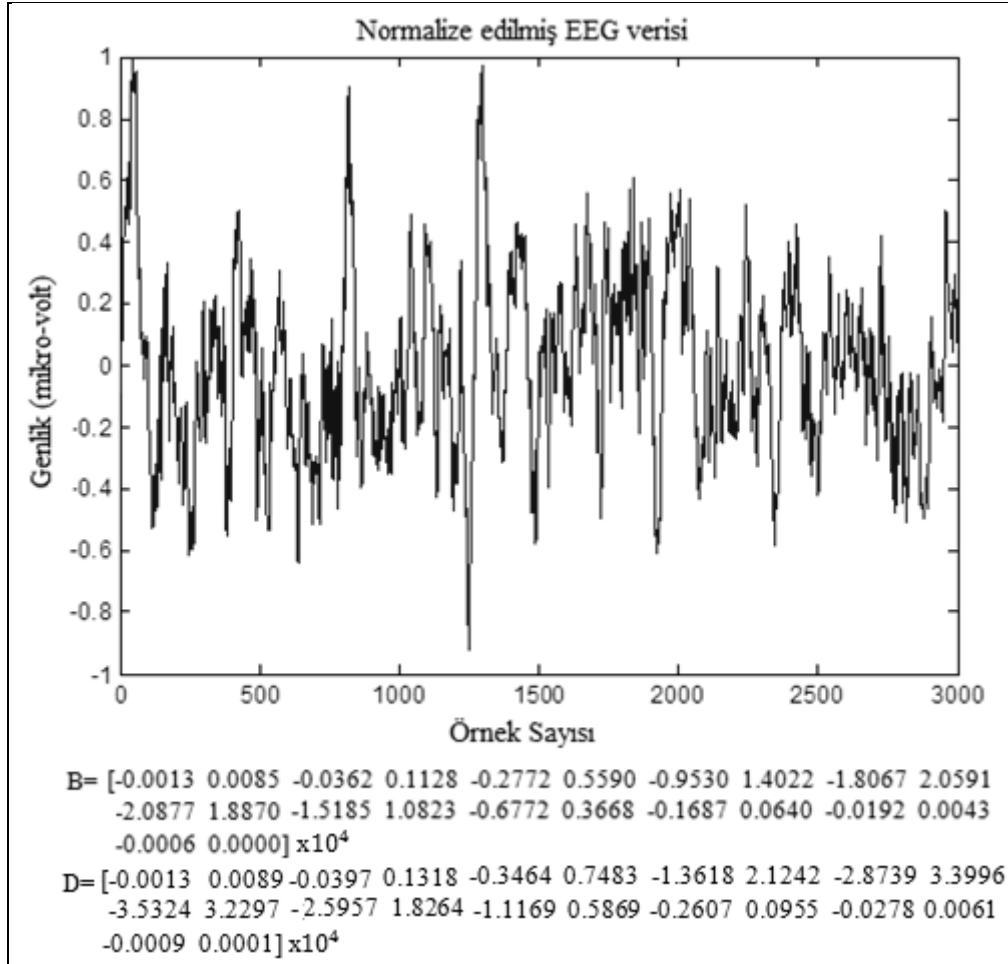
Her kişi için yapılan AIC kriteri ve deneysel tekrarlar sonucunda ideal model derecesinin 22, model algoritmasının ise 'burg' olarak seçilmesi sınıflandırma sonuçlarını olumlu etkilediğinden uygun görülmüştür. Şekil 30-31 ile bir bireye ait birbirinden farklı uyku evrelerinin 30 saniyelik normalize edilmiş EEG verileri görülmektedir. Şekillerin alt kısımlarında bu verilere AR model uygulanması sonucunda elde edilen katsayılar vektör şeklinde verilmiştir. Buradaki önemli nokta farklı evrelere ait farklı EEG verilerinin AR katsayı vektörlerinin de farklı olmasıdır. Aynı zamanda farklı filtre çıkışlarındaki işaretlere uygulanan AR model sonucunda da farklı AR katsayıları elde edilmektedir.



Şekil 30. Normalize edilmiş uyanıklık evresi EEG işaretine ait 30 saniyelik epok örneği

Şekil 30'daki A vektörü, normalize edilmiş uyanıklık EEG verisinin teta filtresinden geçirilmesinin ardından elde edilen çıkışa AR modelin uygulanması sonucunda bulunan AR katsayıları vektörüdür. C vektörü ise yine aynı şekilde uyanıklık evresine ait EEG verisinin alfa filtresinden geçirilmesi sonucunda elde edilen çıkışa AR modelinin uygulanması sonucunda bulunan AR katsayıları vektörüdür.

AR katsayılarıdaki farklılıkları göstermek adına aynı işlemler NREM3 evresine ait olan EEG verilerine de uygulanarak teta ve alfa filtre çıkışlarından sırasıyla B ve D AR katsayı vektörleri şekil 31'deki gibi elde edilmiştir.



Şekil 31. Normalize edilmiş NREM3 evresi EEG işaretine ait 30 saniyelik epok örneği

Aynı işlemler yedi farklı kişi için, üç farklı filtre çıkışına (teta, alfa, beta) AR modelin uygulanması ile tekrarlandı. Sonuç olarak algoritmanın bu kısmında her filtre çıkışı için uyanıklık ve uykunun beş alt evresine ait (NREM1, NREM2, NREM3, NREM4, REM) AR katsayıları (1x22) boyutlu vektörler şeklinde elde edilmiştir.

AR yöntemi kullanılarak bütün kişilere ait her bir band aralığı için elde edilen uyku ve uyanıklık dönemine ait öznelik vektörlerinden 60 tanesi eğitim, 60 tanesi de test kümesi olmak üzere eşit sayıda vektöre sahip iki küme oluşturuldu. Her iki küme verileri içerisinde uyku dönemine ait olan NREM1, NREM2, NREM3, NREM4 ve REM verileri,

uyku verilerini oluşturup '1'etikeri ile etiketlenirirken, uyanıklık dönemine ait olan W verileri ise uyanıklığı temsil eden '0'etikeri ile etiketlenirildiler. Bu işlem gerçekleştirilirken uyku ve uyanıklık verilerine eşit muamele yapmak maksadıyla yukarıda da belirtildiği gibi her iki sınıf için de eşit sayıda veri kullanıldı. Bununla birlikte uyku sınıfını oluşturan her bir alt evre için de (NREM1, NREM2, NREM3, NREM4 ve REM) yine eşit sayıda veri kullanılmıştır. Eğitim ve test veri sayılarının 60 seçilmesinin sebebi ise; yukarıda bahsedildiği gibi her evreye ve her kişiye eşit muamele yapmak adına veri tabanındaki her bir bireye ait uyku ve uyanıklık evrelerinin epok sayıları incelenmiş ve oluşturulacak olan her sınıf için minimum epok sayısına sahip kişiye göre hareket edilmiştir. Veri tabanındaki bütün bireylerin tüm evrelerine ait epok sayıları tablo 3 ile gösterilmiştir.

Tablo 3. Veri tabanındaki her bireyin tüm uyku ve uyanıklık evrelerine ait epok sayıları

	Kişi1	Kişi2	Kişi3	Kişi4	Kişi5	Kişi6	Kişi7
W	(70x3000)	(1824x3000)	(1885x3000)	(2014x3000)	(60x3000)	(128x3000)	(75x3000)
NREM1	(34x3000)	(92x3000)	(59x3000)	(18x3000)	(89x3000)	(121x3000)	(74x3000)
NREM2	(452x3000)	(660x3000)	(373x3000)	(396x3000)	(384x3000)	(396x3000)	(353x3000)
NREM3	(120x3000)	(80x3000)	(94x3000)	(90x3000)	(83x3000)	(53x3000)	(127x3000)
NREM4	(83x3000)	(16x3000)	(203x3000)	(21x3000)	(20x3000)	(127x3000)	(157x3000)
REM	(267x3000)	(176x3000)	(215x3000)	(151x3000)	(216x3000)	(226x3000)	(159x3000)

Bu tablodan da anlaşıldığı üzere 5. Kişinin uyanıklık evresi epok sayısı 60'dır. Bu sebeple uyanıklık evresine ait epok sayıları 60, uyku evresine ait epok sayıları da her bir alt evreden (NREM1, NREM2, NREM3, NREM4 ve REM) 12 epok seçilmek üzere toplam 60 epoktur. Yedi farklı kişiye ait, üç farklı filtreden (teta, alfa, beta) ayrı ayrı geçirilmiş verilerden her filtre çıkışı için elde edilen ve eğitim amaçlı kullanılan uyku- uyanıklık öz nitelik veri boyutları ise tablo 4 ile test amaçlı öz nitelik veri boyutları da tablo 5 ile gösterilmiştir.

Tablo 4. Bir bireye ait eğitim amaçlı oluşturulan öz niteliklerin veri boyutları

Band	Durum	Kişi1	Kişi2	Kişi3	Kişi4	Kişi5	Kişi6	Kişi7
Alfa	uyanıklık	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
	uyku	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
Teta	uyanıklık	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
	uyku	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
Beta	uyanıklık	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
	uyku	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)

Tablo 5. Bir bireye ait test amaçlı oluşturulan özniteliklerin veri boyutları

Band	Durum	Kişi1	Kişi2	Kişi3	Kişi4	Kişi5	Kişi6	Kişi7
Alfa	uyanıklık	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
	uyku	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
Teta	uyanıklık	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
	uyku	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
Beta	uyanıklık	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)
	uyku	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)	(30x22)

Çalışmada kullanılacak olan üç sınıflandırma tekniği için de oluşturulan öznitelik vektörlerinin oluşum aşaması bu kısma kadar anlatıldığı şekildedir. Bu aşamadan sonra farklı sınıflandırma teknikleri ile oluşturulan öznitelik vektörleri uyku ve uyanıklık olarak sınıflandırılıp yapılan sınıflandırmanın başarı sonuçları karşılaştırılacaktır.

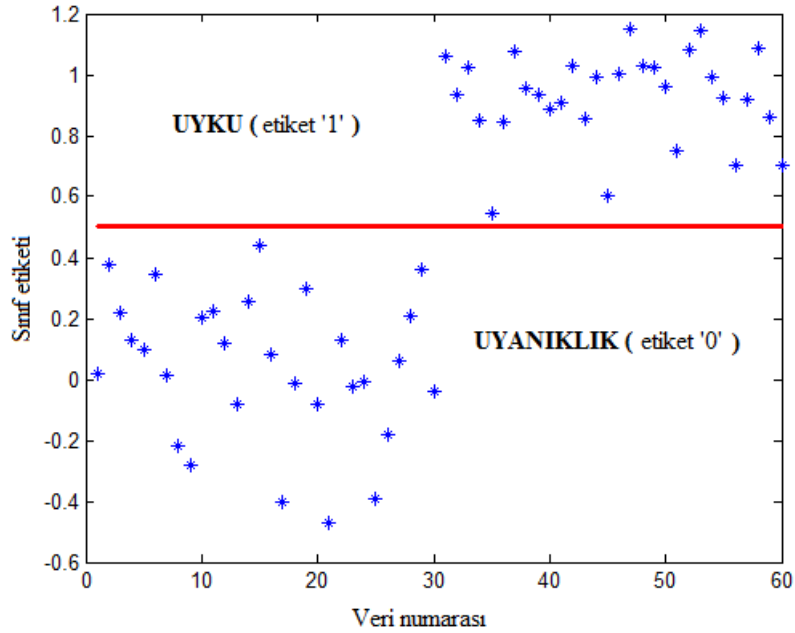
### 3.2. Tahmin Etme ve Sınıflandırmaya Ait Bulgular

#### 3.2.1. Kısmi En Küçük Kareler Regresyon Yöntemine Ait Bulgular

Elde edilen öznitelik verilerinden eğitim amaçlı oluşturulan kümede, her band için var olan özniteliklere KEKKR yöntemi uygulanarak sınıf tahminini gerçekleştirmek için gerekli olan beta ( $\beta_{KEKK}$ ) vektörü elde edilir. Bu işlem için daha önceki bölümlerde de bahsedilen SIMPLS algoritması kullanıldı. Bu parametrenin seçimini güvenilirliği yüksek olacak şekilde sağlamak amacıyla algoritmada çapraz doğrulama (cross validation) yöntemlerinden olan KFOLD metodu kullanıldı. Böylece yüksek doğrulukta sınıflandırma tahminini sağlayan vektör seçimi gerçekleştirilecek ve elde edilen bu beta vektörü test verilerinin sınıf tahmininde kullanılacaktır.

KFOLD yönteminde eğitim verileri seçilen ‘fold’ değeri kadar alt kümeye bölünür ve her bir alt küme sırasıyla ‘onaylama’ kümesi olarak atanırken geriye kalan onaylama kümesi olmayan kümeler ise ‘alt eğitim’ kümesi olarak seçilirler. Böylece tüm eğitim alt kümeleri ‘onaylama’ ve ‘alt eğitim’ kümesi olacak şekilde atanmış olur. Bu çalışmada ‘fold’ değeri 10 olacak şekilde atanmıştır.

Her bir ‘alt eğitim’ verisi kullanılarak elde edilen  $\beta_{KEKK}$  ile ‘onaylama’ verilerinin çarpılması sonucunda elde edilen sınıf etiketlerinin gerçek sınıf etiketleri olan ‘0’ uyanıklık ve ‘1’ uyku değerleri ile karşılaştırılabilmesi için, tahminlenen değerler ‘0.5’ değeri dahil ‘0.5’ değerinden büyük ise ‘1’ sınıfına, küçük ise ‘0’ sınıfına dahil edildi. Kullanılan bu yöntem şekil 32’de 60 veri değerinden oluşan bir eğitim örneği için şekilsel olarak ifade edildi. Şekilden de anlaşılacağı gibi 0.5 etiketi kırmızı çizgi ile gösterilmiştir. Bu çizginin üzerinde kalan veriler uyku etiketi ile altında kalan veriler ise uyanıklık etiketi ile sınıflandırılacaktır.



Şekil 32. Belirlenen kurala göre verilerin etiketlenmesi

Bulunan sınıf tahmin değerleri gerçek değerler ile karşılaştırılıp, denklem (107) ile ifade edildiği gibi doğru tahminlerin ( $Y_{doğru\ tahmin}$ ) bütün tahmin değerlerine ( $Y$ ) oranının 100 ile çarpılması sonucunda sınıflandırma başarısı yüzdesel olarak hesaplanır.

$$\text{Başarı (\%)} = \frac{Y_{\text{doğru tahmin}}}{Y} \times 100 \quad (107)$$

Böylece 10 farklı  $\beta_{KEKK}$  için 10 farklı yüzdesel başarı değeri elde edilir. Bu yüzdesel sonuçlardan en yüksek yüzdeyi veren  $\beta_{KEKK}$  aranan  $\beta_{KEKK}$  olarak kabul edilir. Eğer bu 10 yüzdesel değer arasından aynı maksimum değere sahip birkaç tane yüzdesel başarı değeri mevcut ise bu değerlere karşılık gelen  $\beta_{KEKK}$  vektörlerinin ortalaması alınarak istenen  $\beta_{KEKK}$  elde edilir.

KFOLD yöntemi ile elde edilen  $\beta_{KEKK}$  vektörü eğitim setindeki üç farklı band için de aynı şekilde elde edilir. Sonuç itibariyle üç band için üç farklı  $\beta_{KEKK}$  bulunur. Bir bireye ait eğitim verilerinden elde edilen bu üç ayrı band için bulunan  $\beta_{KEKK}$  vektörleri aynı bandaki test verileri ile çarpılıp, şekil 32 ile ifade edildiği gibi yine '0.5' kuralına göre uyku ve uyanıklık etiketi ile sınıflandırılır. Algoritmadaki bu işlem sonrasında her birey için tablo 6 ile gösterilen sınıflar elde edilir.

Tablo 6. Bir bireye ait her band için tahminlenen sınıfların boyutları

Teta Band için	Alfa band için	Beta Band için
(60x1) boyutlu sınıf tahmini	(60x1) boyutlu sınıf tahmini	(60x1) boyutlu sınıf tahmini

Her bir band için gerçekleştirilen sınıf tahminleri sonucunda tek bir tahmin değeri elde edebilmek için bir kriter belirlendi. Bu kritere göre; en az iki banda ait sınıf tahmin sonuçları birbirleriyle aynı ise veri o sınıf etiketi ile etiketlenecektir. Bu kritere göre bütün olası durumlar için tahmin edilebilecek sınıf etiketleri Tablo 7 ile gösterilmiştir.

Tablo 7. Belirlenen kritere göre olası durumlar ve tahmin edilen etiketler

	Teta band sınıf etiketleri	Alfa band sınıf etiketleri	Beta Band sınıf etiketleri	Kritere göre tahminlenen edilen etiket
Sınıf Etiketleri	0	0	0	0
	0	0	1	0
	0	1	0	0
	0	1	1	1
	1	0	0	0
	1	0	1	1
	1	1	0	1
	1	1	1	1

Bu test işlemi sonrasında EEG verilerinin uyku ya da uyanıklık durumunu ifade etmesine dair sınıf tahminleri gerçek sonuçlar ile karşılaştırılıp denklem (106)'a göre doğruluk değerleri yüzdesel olarak hesaplatıldı.

Sınıflandırma sonrasında elde edilen sonuçların güvenilirliğini daha iyi analiz etmek için bu aşamaya kadar olan işlemler, her evreye ait EEG verileri rastgele olarak karıştırılarak 15 ayrı kez tekrarlanmıştır. Böylelikle tablo 4-5 ile gösterilen veriler 15 farklı kez elde edilmiştir. Sonuç itibariyle 15 ayrı doğruluk yüzdesi elde edildi. Bu yüzdesel doğruluk değerlerinin ortalaması ve standart sapması algoritma ile hesaplatılarak sonuçlar bir tablo halinde tablo 8 ile sunuldu. Bu tablo 60 eğitim ve 60 test öz nitelik kümesi kullanılarak oluşturulmuştur. Her bir eğitim ve test öz nitelik vektörünün boyutu ise 22'dir.

Tablo 8. KEKKR yöntemi sınıflandırma başarı sonuçları

Kişi No.	Ortalama (ort.)	Standart sapma (std.)
Kişi 1	% 91	4.8
Kişi 2	% 96.4	3.2
Kişi 3	% 94.8	3.4
Kişi 4	% 96.8	2.5
Kişi 5	% 93.5	4.6
Kişi 6	% 95.2	2.7
Kişi 7	% 92.4	4.9

Tablodaki sonuçlar incelendiğinde 30 saniyelik epoklar kullanılarak oluşturulan öz nitelik vektörlerine KEKKR yöntemi uygulanarak elde edilen sınıflandırma başarı yüzde değerlerinin yüksek oldukları görülmektedir. 30 saniyelik epoklar kullanılarak çalıştırılan algoritma ayrıca 15, 10 ve 5 saniyelik epoklar kullanılarak da ayrı ayrı tekrarlanmış ve elde edilen sonuçlar tablo 9'da verilmiştir. Epok süresi azaldıkça veri tabanımızda daha fazla öz nitelik vektörü oluşacağından 15, 10 ve 5 saniyelik epoklar kullanılarak tekrarlanan algoritmada eğitim ve test verilerinin sayıları da buna paralel olarak artmıştır. Bu durum elde edilen sonuçların güvenilirliğini daha da arttırmaktadır. Bütün epok süreleri için kullanılan eğitim ve test veri sayıları tabloda belirtilmiştir.



Tablo 9. Farklı epok uzunlukları kullanılarak elde edilen KEKKR yöntemi sınıflandırma başarı sonuçları

	30 saniye		15 saniye		10 saniye		5 saniye	
	60 eğitim-60 test		100 eğitim-100 test		100 eğitim-100 test		300 eğitim-300 test	
	ort.	std	ort.	std.	ort.	std.	ort.	std.
Kişi 1	% 91	4.8	% 89.9	3.8	% 91.6	3.9	% 90.6	2.1
Kişi 2	% 96.4	3.2	% 95.4	1.8	% 95.4	2.5	% 93.8	1.2
Kişi 3	% 94.8	3.4	% 94.7	2.8	% 92	2.7	% 91.9	0.9
Kişi 4	% 96.8	2.5	% 96.8	1.7	% 96.5	1.4	% 96.5	0.6
Kişi 5	% 93.5	4.6	% 92.7	2.7	% 91.4	3.1	% 90.4	2.4
Kişi 6	% 95.2	2.7	% 94.06	2.1	% 92.5	2.6	% 91.5	1.9
Kişi 7	% 92.4	4.9	% 88.2	3.8	% 86.9	4.6	% 86.02	2.5

### 3.2.2. k-En Yakın Komşuluk Yöntemine Ait Bulgular

Çalışmanın bu kısmında, daha önceki bölümde KEKKR yöntemi kullanılarak yapılan sınıflandırma işlemi k-NN yöntemi kullanılarak gerçekleştirildi. 5, 10, 15 ve 30 saniye uzunluğundaki epoklar için ayrı ayrı elde edilen sınıflandırma başarı sonuçları bölümün sonunda bir tablo halinde sunuldu.

K-NN yöntemi kullanılarak yapılan sınıflandırmada tablo 3 ve tablo 4’de boyutları verilen eğitim ve test kümeleri tekrar aynı şekilde oluşturuldu. Sonraki aşamada ise en iyi k değerinin saptanması için KEKKR sınıflandırma yönteminde olduğu gibi KFOLD yöntemi kullanıldı. Optimum k değeri sınavının yapılması için maksimum k değeri 25, fold değeri ise 10 olarak atandı. Her band için var olan eğitim veri seti belirlenen fold sayısı kadar alt parçaya ayrılıp ‘onaylama’ ve ‘alt eğitim’ kümeleri oluşturuldu. Her bir ‘onaylama’ kümesi için ‘alt eğitim’ kümesine olan uzaklıklar denklem (89)’daki Öklid mesafe kuralına göre hesaplandı. Bulunan uzaklık değerleri küçükten büyüğe doğru sıralanıp en küçük olan k tane mesafe belirlendi. Belirlenen bu k en yakın mesafedeki komşuların sınıf etiketleri bulunup sayıca değeri baskın olan sınıf etiketi örüntünün sınıfı olarak atandı. Elde edilen sınıf etiketleri gerçek sınıf etiketleri ile karşılaştırılıp başarı oranı yüzdesel olarak belirlendi.

Fold sayısı 10 olduğundan ötürü bu işlem 10 ayrı ‘onaylama’ kümesi için tekrarlanıp 10 ayrı başarı yüzdesi elde edilir. Bu değerlerden en yüksek başarı yüzdesi, yani en az hata ile doğru sınıflandırma tahmini, o aşamada kullanılan k değeri için geçerli olan başarı yüzdesi olarak kaydedilir. Sınavı gerçekleştirilecek maksimum k değeri daha önceden 25 olarak belirlendiğinden ötürü bu işlem döngüsel olarak 1’den 25’e kadar her k değeri

için tekrarlanır ve sonuçta her k değeri için bir başarı yüzdesi elde edilir bu değerlerden maksimum doğruluk yüzdesini sağlayan k değeri aranan k değeri olarak atanır.

Eğitim aşamasında k değerinin elde edilmesinin ardından test aşamasına geçilir. Üç ayrı banda ait daha önceden tespit edilen test verileri kullanılarak üç bandın sınıf tahminleri k-NN yöntemi ile tekrar ayrı ayrı elde edilir. Sonraki aşamada ise KEKKR yönteminde kullanılan ve tablo 7 ile özetlenen kriter kullanılarak bu üç band için yapılan sınıf tahminlerinden en az iki band tahmin sonuçları birbirleri ile aynı ise test verisi o sınıf etiketiyle etiketlenir.

KEKKR yöntemiyle yapılan sınıflandırma işlemindeki gibi k-NN sınıflandırma yönteminde de her evreye ait EEG verileri rastgele olarak karıştırılarak aynı işlemler 15 ayrı kez tekrarlanmıştır. Sonuç itibarıyla k-NN sınıflandırma yönteminde de 15 ayrı doğruluk yüzdesi elde edilmiştir. Bu yüzdesel doğruluk değerlerinin ortalaması ve standart sapması oluşturulan algoritma ile hesaplatılarak elde edilen sonuçlar tablo 10 ile gösterilmiştir. Sınıflandırma yöntemlerinin başarılarını kıyaslamak adına bu tabloda da 5, 10, 15 ve 5 saniye uzunluğundaki EEG verileri kullanılarak elde edilen başarı yüzdeleri ayrı ayrı verilmiştir.

Tablo 10. Farklı epok uzunlukları kullanılarak elde edilen k-NN yöntemi sınıflandırma başarı sonuçları

	30 saniye		15 saniye		10 saniye		5 saniye	
	(60x22) eğitim (60x22) test		100 eğitim-100 test		100 eğitim-100 test		300 eğitim-300 test	
	ort.	std.	ort.	std.	ort.	std.	ort.	std.
Kişi 1	% 84.1	3.4	% 86.4	2.2	% 86.0	2.9	% 85.3	2.3
Kişi 2	% 96	2.8	% 95.4	2.3	% 94.8	1.6	% 93.2	1.3
Kişi 3	% 94.8	2.8	% 93.5	2.6	% 89.8	3.2	% 91.8	1.09
Kişi 4	% 94.8	2.7	% 94.9	1.7	% 95.3	2.3	% 96.4	1.04
Kişi 5	% 84.6	3.2	% 84.6	3.3	% 83.8	3.7	% 94.4	1.9
Kişi 6	% 91.1	3.9	% 90.6	2.09	% 88.8	2.1	% 86.9	1.6
Kişi 7	% 84.1	3.4	% 85.2	2.8	% 80.6	2.6	% 79.6	2.3

### 3.2.3. Bayes Yöntemine Ait Bulgular

Bu bölümde ise son olarak Bayes yöntemi kullanılarak 5, 10, 15 ve 30 saniye uzunluğundaki epoklar için ayrı ayrı sınıflandırma gerçekleştirilmiştir. Sınıflandırma sonucunda elde edilen başarı oranları bir tablo halinde bölümün sonunda sunulmuştur.

Diğer sınıflandırma yöntemlerinde olduğu gibi bayes sınıflandırma yönteminde de tablo 4 ve tablo 5’de boyutları verilen eğitim ve test kümeleri tekrar aynı şekilde oluşturuldu. Oluşturulan algoritma ile eğitim verilerinin maksimum verimlilikte sınıflandırılmasını gerçekleştiren diskriminant fonksiyonu elde edildi.

Algoritma koşturulmadan önce uyku ve uyanıklık verilerinin olasılıkları 0.5 olarak atandı. Çoklu normal dağılım için olasılık yoğunluk fonksiyonu denklem (108) ile tanımlandığından, diskriminant fonksiyonunu hesaplamak için (108)’in her iki tarafının  $\ln$ ’i alınarak denklem (109) ile verilen diskriminant fonksiyon formülü kullanıldı

$$p(\underline{x}|w_i) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det(\Sigma)}} e^{(-0.5(\underline{x}-\mu)' \Sigma^{-1}(\underline{x}-\mu))} \quad (108)$$

$$g_i(x) = -\frac{1}{2}(\underline{x} - \mu_i)' \Sigma_i^{-1}(\underline{x} - \mu_i) - \frac{K}{2} \ln 2\pi - \frac{1}{2} \ln |\Sigma_i| + \ln P(w_i) \quad (109)$$

Algoritmada ilk olarak eğitim verilerinin 10 ayrı kez karıştırılması ile denklem (109)’daki diskriminant fonksiyonu iki sınıf için de 10 farklı şekilde elde edildi. Her iki sınıf için elde edilen bu 10 farklı diskriminant fonksiyonunun ortalamaları ayrı ayrı alınarak iki sınıf içinde tek bir diskriminant fonksiyon denklemi bulundu. Bu denklemlerde test verileri yerine konularak test verilerinin hangi sınıfa ait olduğuna dair tahminler gerçekleştirildi. Bu işlem diğer sınıflandırma yöntemlerinde olduğu gibi üç band için de ayrı ayrı uygulandı ve sonrasında üç band için de elde edilen sınıf tahminleri diğer yöntemlerde de kullanılan tablo 7’deki kritere göre değerlendirilerek bu üç band için yapılan sınıf tahminlerinden en az iki band tahmin sonuçları birbirleri ile aynı ise test verisi o sınıf etiketiyle etiketlendirildi. Test verileri için bulunan bu sınıf tahminleri gerçek sınıf etiketleri ile karşılaştırılarak denklem (107) ile sınıflandırma başarısı elde edildi.

Tüm bu işlemler yine diğer yöntemlerde olduğu gibi her evreye ait EEG verileri rastgele olarak karıştırılarak 15 ayrı kez tekrarlanmıştır. Sonuç itibariyle bayes sınıflandırma yönteminde de 15 ayrı başarı yüzdesi elde edilmiştir. Bu yüzdesel başarı değerlerinin ortalaması ve standart sapması hesaplatılarak elde edilen sonuçlar tablo 11 ile gösterilmiştir. Aynı şekilde sınıflandırma yöntemlerinin başarılarını kıyaslamak için bu tabloda da 5, 10, 15 ve 5 saniye uzunluğundaki EEG verileri kullanılarak elde edilen başarı yüzdeleri ayrı ayrı verilmiştir.

Tablo 11. Farklı epok uzunlukları kullanılarak elde edilen bayes yöntemi sınıflandırma başarı sonuçları

	30 saniye		15 saniye		10 saniye		5 saniye	
	60 eğitim-60 test		100 eğitim-100 test		100 eğitim-100 test		300 eğitim-300 test	
	ort.	std.	ort.	std.	ort.	std.	ort.	std.
Kişi 1	% 80.5	2.9	% 83.8	3.0	% 83.8	4.5	% 84.7	3.1
Kişi 2	% 81.3	11.3	% 86.8	9.2	% 86.5	8.7	% 89.8	2.3
Kişi 3	% 72.1	19.3	% 88.3	6.0	% 84.5	9.3	% 88.9	6.6
Kişi 4	% 64.7	14.9	% 86.2	14.4	% 84.4	13.9	% 90.5	10.1
Kişi 5	% 78.5	7.9	% 85.8	3.6	% 85.8	3.6	% 85.8	6.1
Kişi 6	% 81.7	4.2	% 84.4	5.6	% 83.4	4.5	% 80.1	5.3
Kişi 7	% 80.5	2.9	% 82.9	3.9	% 79.8	5.2	% 81.7	3.16

Uyku EEG'sinin üç farklı sınıflandırma yöntemi kullanılarak gerçekleştirilen sınıflandırma işleminde elde edilen başarı ortalamaları ve standart sapma değerleri tablo 9-11 ile verilmiştir. Bu tablolar incelendiğinde genel anlamda sınıflandırma yüzdesel başarı ortalamalarının 7 numaralı kişi hariç KEKKR yöntemiyle yapılan sınıflandırmada % 90'ın üzerinde, k-NN yöntemiyle yapılan sınıflandırmada KEKKR yöntemine göre biraz daha düşük, bayes yönteminde ise yine % 80' ler seviyesinde olduğu görülmektedir. 7 numaralı kişide ise her üç yöntemde de başarı oranı genel anlamda diğerlerine göre düşük çıkmıştır. Bunun sebebinin daha önce materyal bölümünde de bahsedilen uykuya dalma güçlüğü yaşayan birey olmasından ötürü olduğu düşünülmektedir. Standart sapma değerleri incelenecek olursa KEKKR ve k-NN yönteminde 5 saniye uzunluğundaki verilere ait standart sapma değerlerinin diğer gruplara göre düşüş gösterdiği görülmektedir. Bunun anlamı farklı eğitim setleri için hesaplanan sınıflandırma başarı değerleri arasındaki farklılığın 5 saniyelik veri grupları için az olmasıdır. Çünkü 5 saniyelik veri gruplarında daha çok veri eğitime tabi tutulmuştur ve bu durum da sınıflandırma başarı oranını arttırmaktadır. Bayes yöntemiyle yapılan sınıflandırmada ise bu duruma ait net bir sonuç çıkarılamamaktadır. Ayrıca 5 saniyelik veri uzunluğuna sahip kümelerin diğer veri kümelerine göre daha güvenilir sonuçlar vermesi gerçel zaman uygulamalı işlemlerde zaman ve hız açısından daha olumlu bir neticedir.

Tüm bu sonuçlar dışında bir diğer önemli nokta ise her üç sınıfın eğitim ve test aşamalarında harcadıkları zamandır. Gerçek zamanlı uygulamalarda uyku ile uyanıklığın sınıflandırılmasında olabildiğince kısa sürede bu ayrımın yapılabilmesi zaman ve hız tasarrufu açısından önemli bir noktadır. Bu çalışmada her üç sınıflandırma yönteminde de

öz niteliklerin elde edilmesine kadar geçen süreç aynıdır. Normalizasyon işlemi uygulanmış ve filtrelenmiş 30, 15, 10 ve 5 saniye uzunluğundaki bir epoktan 22 AR katsayısının elde edilmesi için için geçen süre yaklaşık olarak  $\sim 0.1-0.2$  saniyedir. Epok uzunluklarının AR katsayısının hesap süresindeki etkisi ise yaklaşık olarak 0.01-0.05 saniye düzeyindedir ve en kısa hesap süresi 5 saniye uzunluğundaki epoklara aittir. Bu yüzden özellikle gerçek zamanlı uygulamalarda ve çok sayıda verinin eğitilmesinde zaman ve hız açısından 5 saniyelik epokların kullanılması daha avantajlıdır.

Öz niteliklerin elde edilmesinin ardından sınıflandırma aşamasında KEKKR, k-NN ve bayes yöntemi kullanılmıştır. Her üç yöntem için (60x22) boyutlu eğitim ve (60x22) boyutlu test verilerinden oluşan öz nitelik kümesinin açıklamalarda bahsedildiği şekilde (15 ayrı veri kümesi ve her küme için 3 band) eğitilmesi ve sonrasındaki test aşamasında harcanan süreler tablo 12 ile verilmiştir.

Tablo 12. Her üç sınıflandırma yönteminde eğitim ve test aşamasında harcanan süreler

	KEKKR	k-NN	BAYES
EĞİTİM	$\sim 2.3$ saniye	$\sim 220$ saniye	$\sim 3.5$ saniye
TEST	$\sim 0.4$ saniye	$\sim 6.75$ saniye	$\sim 1.2$ saniye

Bu tablodaki test süreleri 15 ayrı veri kümesine ait (1x22) boyutlu 60 farklı öz nitelik vektörü içindir. Tek bir örnek için harcanan süre ise bu sürelerin  $1/60$ 'dir.

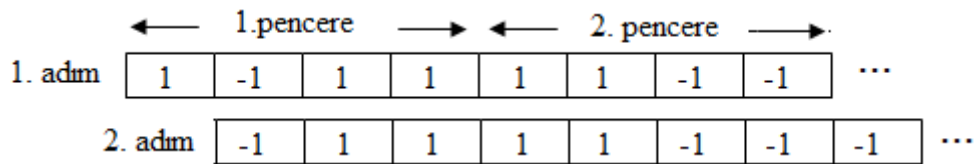
Tüm bu sonuçlar incelendiğinde k-NN ve bayes yönteminin başarı oranlarının KEKKR yöntemine göre düşük olduğu görülmektedir. Bunun yanında k-NN yönteminin eğitim ve test aşamasında harcanan süre de diğer yöntemlere göre oldukça uzundur. Bunun sebebi eğitim ve test aşamasında bir verinin eğitim kümesinde bulunan tüm örneklerle uzaklığının incelenmesidir. Bu işlem algoritmanın sürecini uzatmaktadır. Bayes yönteminde ise harcanan süre k-NN yöntemine göre çok daha kısa olmasına rağmen KEKKR yönteminde harcanan süreye göre daha uzundur. Ayrıca bayes yöntemi ile elde edilen doğruluk tablosundaki standart sapma değerlerinin diğer yöntemlere göre yüksek olması bu başarı oranlarının bazı test örneklerinde oldukça düşüş gösterdiğini belirtmektedir. Bu da güvenilirlik açısından olumsuz bir sonuçtur. KEKKR yöntemi ise hem başarı ortalamaları hem de standart sapma ve eğitim ile test süre değerleri açısından diğer iki yönteme göre daha güvenilir sonuçlar vermiştir. Bu da KEKKR yönteminin diğer yöntemlere göre EEG'de uyku ve uyanıklık durumunun ayırt edilmesindeki başarısını

ortaya koymaktadır. Gerçek zamanlı uygulamalarda ise KEKKR yönteminin 5 saniye uzunluğundaki epoklar için kullanılması zaman açısından daha iyi sonuçlar verecektir.

### 3.2.4. Yöntemin Performansına İlişkin Bulgular

EEG kayıtlarındaki uyanıklık ile uyku durumunun KEKKR yöntemi kullanılarak başarılı bir şekilde sınıflandırılmasının ardından uyanıklıktan uykuya geçiş anlarının kısa zaman dilimlerinde yakalanabilmesi için bir performans çalışması yapılmıştır. Bu çalışmada KEKKR yönteminin eğitim aşamasında elde edilen  $\beta$  vektörünün bireylere özgü olarak elde edilmesinin ardından her bireye ait kesintisiz sürekli EEG kayıtlarının bu  $\beta$  vektörü ile test edilmesiyle sınıf tahminleri yapılmış ve uyanıklıktan uykuya geçiş anlarını gösteren geçiş grafikleri elde edilmiştir. Bu geçiş grafiklerini elde etmek için uygulanan yöntem ise şu şekildedir;

Uzunluğu 4 birim olacak şekilde oluşturulan iki dikdörtgen pencere peşpeşe ilerleyecek şekilde ve örtüşmeden sınıf tahmini gerçekleştirilen sürekli EEG kayıtları üzerinde adım adım ilerletilir. Her adım boyutu 30 saniyelik bir epok uzunluğunda olup dikdörtgen pencerelerin her bir parçasına bir epok karşılık gelmektedir. Pencerelerin bir birimine karşılık gelen epokun etiketi uyanıklık ise pencerenin o parçası '1' etiketi ile, uyku ise '-1' etiketi ile isimlendirilir. Ardından her bir pencerenin ortalaması hesaplanıp ikiye bölünür (ikiye bölme işlemi işlem sonunda ölçek değerlerinin [-1 1] aralığında olması için yapılmıştır). Bahsedilen durum şematik olarak şekil 33 ile gösterilmiştir.



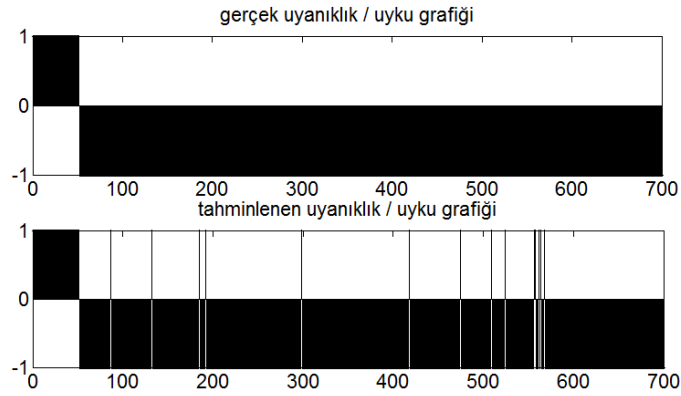
Şekil 33. Performans çalışması için şematik gösterim

Örneğin şekil 33' deki birinci adım için 1. pencerenin ortalama hesabı  $ort1 = 0.5 * (2/4) = 0.25$  şeklinde, 2. pencerenin ortalama hesabı ise  $ort2 = 0.5 * (0/4) = 0$  şeklinde hesaplanır. Bu işlemin ardından her adımda her iki pencere için hesaplanan ortalama değerlerinin farkı ( $ort1 - ort2$ ) hesaplatılıp oluşan indekse göre bu farkların grafiği

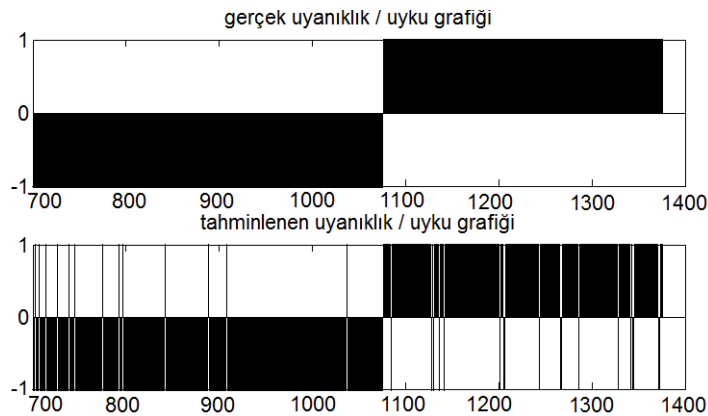
çizilir. Elde edilen grafikte ‘1’ değeri uyanıklıktan uykuya geçiş anlarını, ‘-1’ değeri ise uykudan uyanıklığa geçiş anlarını ifade etmektedir.

Bahsedilen bu performans çalışmasının yedi farklı kişiye ait EEG kayıtlarına uygulanması sonucunda elde edilen en iyi ve en kötü performans sonuçları şekil 34-38 ile verilmiştir.

Şekil 34-36 bir bireye ait 11.4 saatlik EEG kaydının uyanıklık ve uyku olarak etiketlenen evrelerini göstermektedir. Veri boyutu uzun olduğu için grafikler iki şekil ile verilmiştir. Yani şekil 35, şekil 34’ ün devamıdır. Her iki şekildeki ilk grafik bu veriye ait gerçek sınıfları ikinci grafik ise KEKKR yöntemi ile elde edilen sınıf tahminlerini göstermektedir. Bu grafiklerdeki ‘1’ değerleri uyanıklık durumunun, ‘-1’ değerleri ise uyku durumunun etiketleridir. Denklem (107) kullanılarak tahmin edilen bu sınıf etiketlerinin gerçek sınıf etiketleri ile karşılaştırılması sonucunda % 96.6 tahminleme başarı yüzdesi elde edilmiştir.

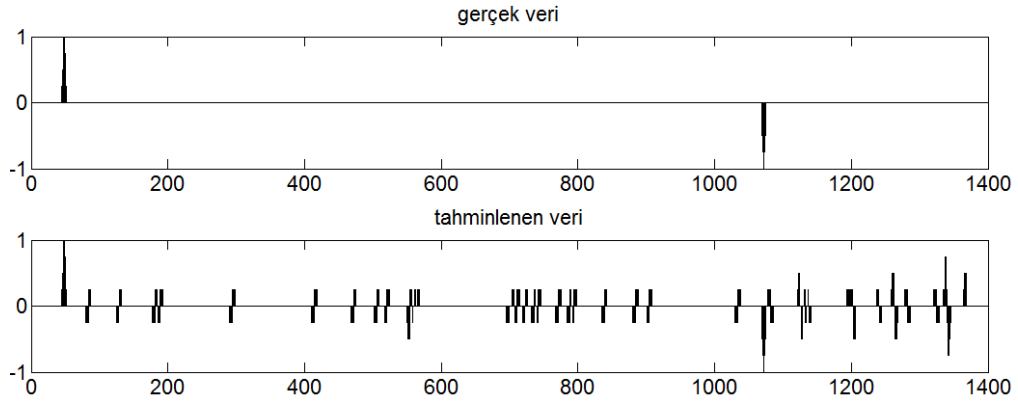


Şekil 34. Gerçek ve tahmin edilen veriler için uyanıklık/uyku grafikleri



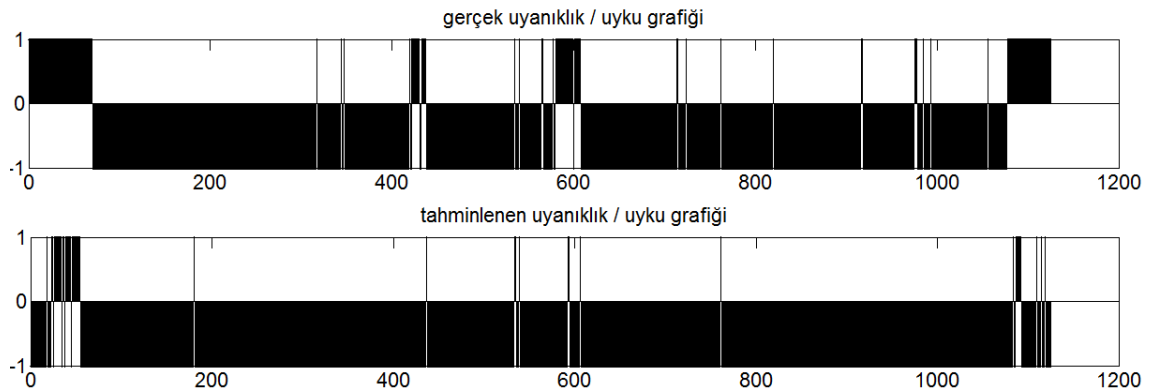
Şekil 35. Gerçek ve tahmin edilen veriler için uyanıklık/uyku grafikleri

Şekil 36 ise aynı bireye yukarıda bahsedilen performans çalışmasının uygulanması sonucunda elde edilmiş olup şekildeki grafikler incelendiğinde uyanıklık/ uyku geçiş anları yukarıda anlatıldığı gibi '1' ve '-1' etiketleri ile gözlemlenebilmektedir. Şekildeki ilk grafik gerçek verideki bu geçiş anlarını ikinci grafik ise tahmin edilen verideki geçiş anlarını göstermektedir. Tahmin edilen verideki geçiş anları incelendiğinde '1' ile ifade edilen uykudan uyanıklığa geçiş anı ve '-1' ile ifade edilen uyanıklıktan uykuya geçiş anlarının gerçek verideki geçiş anları ile örtüştüğü görülmektedir.



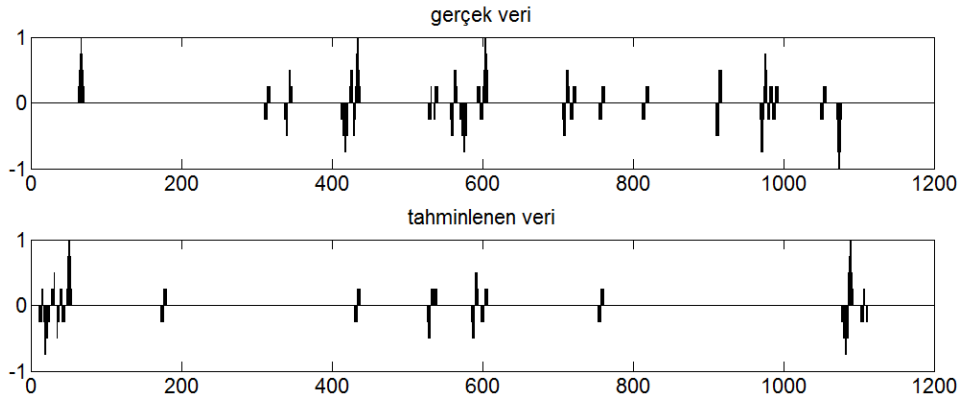
Şekil 36. Uyanıklık-uyku geçiş grafikleri

Aynı incelemeler sonucunda elde edilen en kötü performansı gösteren bireye ait grafikler ise şekil 37-38 ile verilmiştir. Bu bireye ait EEG kayıt uzunluğu 9.3 saat olup tahmin başarı yüzdesi % 87.4' tür. Şekil 38' deki uyanıklık-uyku geçiş grafiğinde gerçek verideki ilk uyanıklıktan uykuya geçiş anının doğru tahmin edildiği fakat diğer anlarda hatalar olduğu görülmektedir.



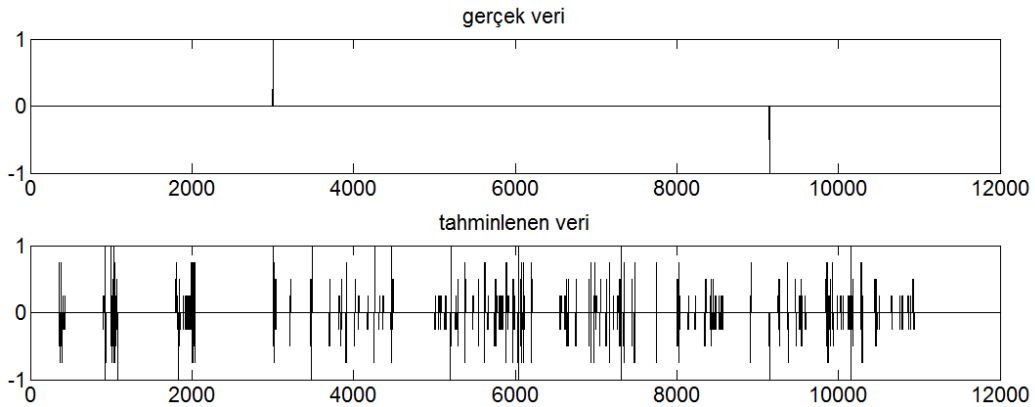
Şekil 37. Gerçek ve tahmin edilen veriler için uyanıklık/uyku grafikleri



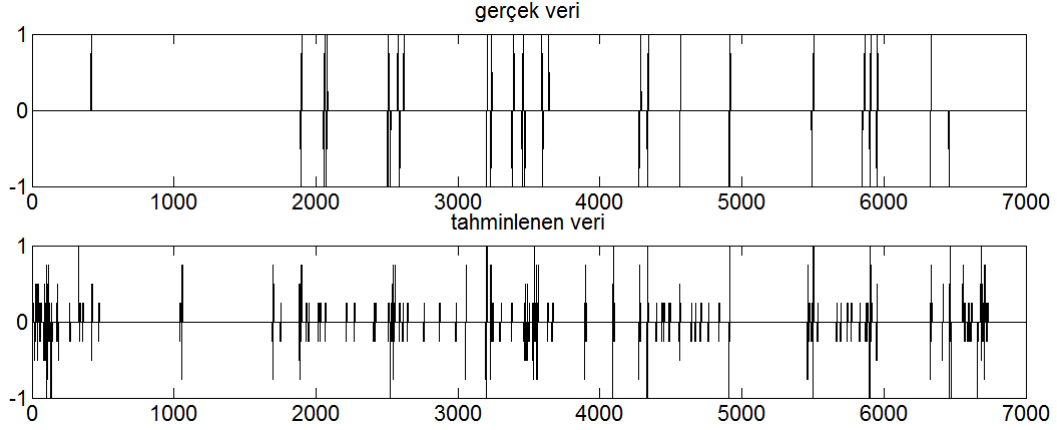


Şekil 38. Uyanıklık-uyku geçiş grafikleri

30 saniyelik epok adımları ile ilerleyerek yapılan bu performans çalışmasında uyanıklıktan uykuya geçiş anları pencere boyutu olan ( $4 * 30 = 120$ ) saniye yani 2 dakikalık süreç gerektirmektedir. Aynı işlemler 5 saniyelik epok adımları kullanılarak gerçekleştirildiğinde ise bir pencere boyutu ( $4*5=20$ ) saniye yani yarım dakikadan kısa bir süreye ineceğinden daha hızlı bir hesaplama gerçekleşecektir. Aynı bireyler için beş saniyelik epok adımları kullanılarak elde edilen uyanıklık-uyku geçiş grafikleri şekil 39 ve şekil 40 ile verilmiştir. Yine aynı şekilde şekil 39 ile yüksek performans gösteren, şekil 40 ile de düşük performans gösteren bireye ait grafikler verilmiştir. Uyanıklık/uyku grafikleri ise 5 saniye epok uzunlukları için toplamda çok uzun boyuta sahip olduklarından ötürü geçişler iyi gözlemlenemeyeceğinden grafik olarak eklenmemiştir. Bu grafikler incelendiğinde uyanıklık-uyku geçiş anlarını ifade eden '1' ve '-1' noktalarının gerçek geçiş noktaları ile karşılaştırıldığında 30 saniyelik epok adımlı performans çalışmasına göre daha iyi sonuçlar verdiği gözlemlenmektedir.



Şekil 39. Uyanıklık-uyku geçiş grafikleri



Şekil 40. Uyanıklık-uyku geçiş grafikleri

Performans örnekleri verilen bu iki kişi dahil tüm 7 kişiye ait sürekli EEG verileri için KEKKR yöntemi kullanılarak yapılan uyanıklık/uyku sınıflandırmasına ait yüzdesel başarı değerleri ise tablo 13 ile verilmiştir. Tabloda 5 saniye ve 30 saniye epok uzunluğundaki verilere ait sonuçlar ayrı ayrı ifade edilmiştir. Yukarıdaki uyanıklık-uyku geçiş grafikleri verilen yüksek ve düşük performans gösteren kişiler 2. ve 3. kişilerdir. Tablo incelendiğinde bu kişilere ait yüzdesel başarı değerlerinden de bu sonuçlar anlaşılmaktadır.

Tablo 13. Sürekli EEG kayıtları için KEKKR yöntemi ile sınıflandırma başarısı

Kişi No.	Başarı (30 saniyelik epok adımı için)	Başarı (5 saniyelik epok adımı için)
Kişi 1	% 93.7	% 95.3
Kişi 2	% 96.6	% 96.5
Kişi 3	% 87.4	% 89.0
Kişi 4	% 97.5	% 96.4
Kişi 5	% 96.0	% 94.4
Kişi 6	% 95.8	% 89.6
Kişi 7	% 90.6	% 91.0

Uygulanan bu performans çalışmasında elde edilen uyanıklık-uyku geçiş grafiklerindeki sayısal değerlerde yapılacak olan incelemeler sonucunda daha farklı istatistiksel sonuçlar çıkarılabileceği düşünülmektedir. Ayrıca kullanılan pencerenin boyutu azaltılarak yapılan bu performans çalışması daha da ileri aşamalara götürülebilir. Sürekli EEG kayıtlarına ait uyku ve uyanıklığın birbirinden ayırt edilmesine ait yüzdesel

başarı sonuçları 7 farklı kişi için tablo 13 ile gösterilmiş olup KEKKR yönteminin sınıflandırmadaki başarısı tekrar ortaya konulmuştur. Bu tablonun önceki KEKKR sınıflandırma yöntemine ait başarı tablosundan (tablo 9) farkı burada kullanılan EEG kayıtlarının kesintisiz sürekli kayıtlar olmasıdır. Yani test edilen veri sayısı çok daha fazladır ve her birey için toplam uyku süreleri farklılık gösterdiğinden ötürü bu sayı her bir birey için de farklıdır. Bunu yanında eğitim aşamasında kullanılan beta vektörü ise aynı beta vektörüdür. Tablo 9 ile karşılaştırıldığında sürekli EEG verileri kullanılarak elde edilen başarı değerlerinde 3. Kişinin yüzdesel değerinde bir parça azalma gözlemlenmiştir. Bunun nedeninin ise sürekli EEG kayıtlarında kayıt sırasında yapılan hatalardan kaynaklandığı düşünülen ve değer olarak sonsuza yakın değer alan hatalı epokların olduğu düşünülmektedir.

#### 4. TARTIŞMA VE ÖNERİLER

Yapılan bu tez çalışmasında EEG kayıtlarında uyku ve uyanıklık durumunun birbirlerinden ayırt edilebilmesi amaçlanmış ve bu kayıtlarda uyku başlangıcı erken olarak yakalanmaya çalışılmıştır. Bu amaçla üç farklı sınıflandırma yöntemi ayrı ayrı denenmiş ve KEKKR yönteminin diğer iki yonteme göre hız, zaman ve başarı yönünden üstünlüğü ortaya konulmuştur. Algoritmanın koşturulması sonucunda elde edilen başarı değerleri tablolardan da anlaşılacağı üzere güvenilir sonuçlar vermiştir.

Çalışmada kullanılan EEG verileri kafatasının tek bir bölgesi olan 'PzOz' bölümünden elde edilen verilerdir. Tek kanalın kullanılmış olması daha önce bu konuyla ilgili yapılan çok kanallı araştırmalara göre üstünlük sağlamaktadır. Bunun yanında 5 saniyelik veriler için iyi sonuçlar vermiş olması zaman açısından gerçek zamanlı uygulamalarda olumlu bir etki oluşturmaktadır.

Physionet veri tabanındaki kayıtların sınırlı olması nedeniyle kullanılan algoritmada eğitim ve test veri sayısı belirli sınırlar içerisinde kalmıştır. Daha fazla EEG verisi kullanılarak çalışmadan daha verimli sonuçlar elde edilebileceği düşünülmektedir. Yapılan bu çalışmanın daha da ileri götürülüp kullanılan sınıf sayısının artırılması ve uykunun diğer alt evrelerinin de uyanıklıktan ayırt edilebilmesi hedeflenilmektedir. Ayrıca algoritmada kullanılan alfa, teta, beta frekans bandlarına bir de derin uyku evresinde ortaya çıkan delta bandının da dahil edilip daha farklı sınıf etiketi belirleme kriterleri kullanılması planlanmaktadır.

Çalışmadaki sınıflandırma tahminlerindeki hatalı sınıflandırmalar incelendiğinde ise hatalı tahminlerin daha çok uyanıklık ile uyku başlangıcı olan NREM1 evresine ait sınıf etiketlerinde olduğu gözlemlenmiştir. Ayrıca uyanıklık EEG'si ile REM dönemine ait EEG kayıtlarında benzerlik olmasının da sonuçları etkilediği ve hatalı tahminlere yol açtığı düşünüldüğünden bu konuda da ayrı olarak araştırma yapılacaktır. Çalışmada KEKKR yöntemi sonucunda elde edilen  $\beta$  vektörü her kişiye göre değişen bir vektör olup ilerleyen çalışmalarda bu vektörün tüm kişiler için tek bir vektör olarak elde edilebilmesi üzerine çalışılacaktır. Elde edilen bu çalışma sonuçlarının klinik uygulamalarda kullanımının yanında uykunun erken teşhisinin sağlanmaya çalışıldığı gerçek zamanlı yorgunluk ikaz sistemlerinin geliştirilmesinde de kullanılabileceği düşünülmektedir.

## 5. KAYNAKLAR

1. Duman, F., O., EEG Sinyallerinin Analizi ile Uyku Durumunun Belirlenmesi, Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2005.
2. Berger, H., Über das Elektrenkephalogramm des Menschen, Archives für Psychiatrie und Nervenkrankheiten, 8 (1926) 527-570.
3. Yazgan, E. ve Korürek, M., Tıp Elektroniği, Birinci Baskı, İTÜ Elektrik-Elektronik Fakültesi Ofset Baskı Atölyesi, İstanbul, 1996.
4. Malmivuo, J. ve Plonsey R., Bioelectromagnetism, Oxford University Press, Newyork, 1995.
5. Susmakova, K., Human Sleep and Sleep EEG, Measurement Science Rewiev, 4 (2004) 59-74.
6. Erdamar, A., Uyku Apnesinin Öngörülmesi ve Dil Kasının Uyarılması için Model Geliştirilmesi, Doktora Tezi, Hecettepe Üniversitesi, Ankara, 2007.
7. <http://uykubozukluđu.uludag.edu.tr/uykulaboratuvarı.htm>, Uyku Laboratuvarında Neler Yapılıyor, 1 Ekim 2012.
8. Bayrak, T., Koçak, O., Erdamar, A., Uyku Bozukluklarının Araştırılmasında Biyomedikal Mühendislik Yaklaşımları. [http://www.emo.org.tr/ekler/0261701767cb4b8\\_ek.pdf?dergi=657](http://www.emo.org.tr/ekler/0261701767cb4b8_ek.pdf?dergi=657), 1 Ekim 2012.
9. Pıçak, R., İsmailoğulları S., Mazıcıoğlu, M. M., Üstünbaş, H. B. ve Aksu, M., Birinci Basamakta Uyku Bozukluklarına Yaklaşım ve Öneriler, Turkish Journal of Family Medicine and Primary Care, 3 (2010) 12-22.
10. Williams, R.L., Karacan, İ. ve Hirsch, C.J., Electroencephalography (EEG) of Human Sleep: Clinical Application, A Willey Biomedical Health Publication, Newyork, 1974.
11. Asyalı, M. H., Berry, R. B., Khoo, M. C. K. ve Altınok A., Determining a Continuous Marker for Sleep Depth, Computers in Biology and Medicine, 37 (2007) 1600-1609.
12. Chapotot, F. ve Becq, G., Toh, Automated Sleep-Wake Staging Combining Robust Feature Extraction, artificial Neural Network Classification and Flexible Decision Rules, International Journal of Adaptive Control and Signal Processing, 24 (2010) 409-423.

13. Krakovska, A. ve Mezeiova, K., Automatic Sleep Scoring: A search for an Optimal Combination of Measures, Artificial Intelligence in Medicine, 53 (2011) 25-33.
14. Khasawneh, N., Jaradat, M. A. K., Fraiwan, L. ve Al-Fandi, M., Adaptive Neuro-Fuzzy Inference System for Automatic Sleep Multistage Level Scoring Employing EEG, EOG and EMG Extracted Features, Applied Artificial Intelligence, 25 (2011) 163-179.
15. Adnane, M. S., Jiang, Z. Y. ve Yan, Z., Sleep-Wake Stages Classification and Sleep Efficiency Estimation Using Single-Lead Electrocardiogram, Expert Systems with Applications, 39 (2012) 1401-1413.
16. U. Acharya, R., Faust, O., Kannathal, N., Chua, T. ve Laxminarayan, Non-Linear Analysis of EEG Signals at Various Sleep Stages, Computer Methods and Programs in Biomedicine, 80 (2005) 37-45.
17. Görür, D., Halıcı, U., Aydın, H., Ongun, G., Ozgen, F. ve Leblebicioğlu, K., Sleep Spindles Detection Using Short Time Fourier Transform and Neural Networks, International Conference on Neural Networks, 2002, Paris , 2, 1631-1636.
18. Alvarez-Estevez, D. ve Moret-Bonillo, V., Identification of Electroencephalographic Arousals in Multichannel Sleep Recordings, IEEE Transactions on Biomedical Engineering, 58 (2011) 54-63.
19. Jo, H. G., Park, J. Y., Lee, C. K., An, S. K. ve Yoo, S. K., Genetic Fuzzy Classifier for Sleep Stage Identification, Computers in Biology and Medicine, 40 (2010) 629-634.
20. Güneş, S., Polat, K. ve Yosunkaya, Ş., Efficient Sleep Stage Recognition System Based on EEG Signal Using K-Means Clustering Based Feature Weighting, Expert Systems With Applications, 37 (2010) 7922-7928.
21. Mariani, S., Manfredini, E., Rosso, V., Grassi, A., Mendez, M. O., Alba, A., Matteucci, M., Parrino, L., Terzano, M. G., Cerutti, S. ve Bianchi, A. M., Efficient Automatic Classifiers for The Detection of a Phases of Cyclic Alternating Pattern in Sleep, Med Biol Comput, 50 (2012) 359-372.
22. Bayram, M., EEG Sinyalleriyle Uyuklama Seviyesinin Modern Yöntemlerle Kestirimi, Yüksek Lisans Tezi, Dicle Üniversitesi, Fen Bilimler Enstitüsü, Diyarbakır, 2003.
23. Batar, H., EEG İşaretlerinin Dalgacık Analiz Yöntemleri Kullanılarak Yapay Sinir Ağları ile Sınıflandırılması, Yüksek Lisans Tezi, Kahramanmaraş Sütçü İmam Üniversitesi, Fen Bilimler Enstitüsü, Kahramanmaraş, 2005.
24. Torun, S. Ö., Uyku EEG'sinde Karşılaşılan İğciklerin(Spindle) Sezimi Üzerine Bir Çalışma, Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2005

25. Akdemir, B., Tahmin Uygulamalarında Performans Geliştirmek İçin Kullanılan Normalizasyon Metotlarına Yeni Bir Yaklaşım, Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya, 2009.
26. Semmlow, J. L., Biosignal ve Biomedical Image Processing, Matlab-Based Applications, 10, CRC Press, New Jersey, 2004.
27. Sevüktekin, M. ve Nargeleçekenler, M., Zaman Serileri Analizi, 1. Baskı, Nobel Basımevi, Ankara, 2005.
28. Akgül, I., Zaman Serilerinin Analizi ve ARIMA Modelleri, 1. Baskı, Der Yayınevi, İstanbul, 2003.
29. Box, G. E., Jenkins, G. M. ve Reinsel, G. C., Time Series Analysis Forecasting and Control, Fourth Edition, John Wiley & Sons, New Jersey, 2008.
30. Haykin, S., Neural Networks: A Comprehensive Foundation, First Edition, Macmillian College, Newyork, 1994.
31. Palaniappan, R., Biological Signal Analysis, First Edition, Ramaswamy Palaniappan & Ventus Publishing, 2010.
32. Kıymık, M. K., Subaşı, A., Dizibüyük, A. ve Özer, M. S., Darbeli Doppler Laminar Kan Akış Sinyal Simülasyonuna STFT ve AR Spektral Analizlerinin Uygulanması, KSÜ J. Science and Engineering, 5, 2 (2002) 14-27.
33. Frank, I. E. ve Friedman, J. H., A Statistical View of Some Chemometrics Regression Tools, Technometrics, 35, 2 (1993) 109-148.
34. Geladi, P. ve Kowalski, B., Partial Least Square Regression: a tutorial, Analytica Chimica Acta, 185, 1986, 1-17.
35. Höskuldsson, A., PLS Regression Methods, Journal of Chemometrics, 2 (1988) 211-228.
36. Phatak, A. ve De Jong, S., The Geometry of Partial Least Squares, Journal of Chemometrics, 11 (1997) 311-338.
37. Ter Braak, C. J. F. ve De Jong S., The Objective Function of Partial Least Squares Regression, Journal Of Chemometrics, 12 (1998) 41-54.
38. Wold, S., Personal Memories of the Early PLS Development, Chemometrics and Intelligent Laboratory Systems, 58, 2 (2001) 83-84.
39. Abdi, H., Partial Least Squares Regression ve Projection on Latent Structure Regression, Wiley Interdisciplinary Reviews: Computational Statistics, 2, 1 (2010) 97-106.

40. Polat, E., Kısmi En Küçük Kareler Regresyonu, Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2009.
41. Trygg, J., Have you ever wondered why PLS sometimes needs more than one component for a single-y vector? <http://www.chemometrics.se/editorial/feb2002.html> 10 Eylül 2012.
42. Bulut, E., Kısmi En Küçük Kareler Regresyon Yöntemi Algoritmalarından Nipals ve PLS-Kernel Algoritmalarının Karşılaştırılması ve Bir Uygulama, Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 24, 2 (2009) 127-138.
43. Hoerl, A. E. ve Kennard, R. W., Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics, 12, 1 (1970) 69-82.
44. Rawlings, J. O., Applied Regression Analysis: a research Tool, Pacific Grove, California, 1988.
45. Naes, T., Isaksson, T., Fearn, T. ve Davies, T., A User-Friendly Guide to Multivariate Calibration and Classification, NIR Publication Chichester, UK, 2003.
46. Naes, T. ve Martens, H., Comparison of Prediction Methods for Multicollinear Data, Communications in Statistics-Simulation and Computation, 14, 3 (1985) 545-576.
47. Koçak, İ., Çok Değişkenli Çoklu Ayarlama Problemlerinde Değişken Seçimi, Doktora Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2006.
48. Martens, H. ve Naes, T., Multivariate Calibration, John Wiley & Sons, Newyork, 1989.
49. Yeniay, Ö., A Comparison of Partial Least Squares Regression with Other Prediction Methods, Hacettepe Journal of Mathematics and Statistics, 31 (2002) 99-111.
50. Lapinsh, M., Development of Methods for the Analysis of Ligand-Protein Interactions. [www.lcb.uu.se/data/LapinshPresentation.ppt](http://www.lcb.uu.se/data/LapinshPresentation.ppt) 15 Ekim 2012.
51. Wold, S., Sjöström, M. ve Eriksson, L., PLS-Regression: A Basic Tool of Chemometrics, Chemometrics and Intelligent Laboratory Systems, 58 (2001) 109-130.
52. Wold, S., Eriksson, L., Trygg, J. ve Kettaneh, N., The PLS Method-Partial Least Squares Projections to Latent Structures-and Its Applications in Industrial RDP (research, development, and production), Compdat, Prague.
53. Lindgren, F. ve Rannar, S., Alternative Partial Least-Squares (PLS) Algorithms, Perspectives in Drug Discovery and Design, 12-14 (1998) 105-113.
54. De Jong, S., SIMPLS: An Alternative Approach to Partial Least Squares Regression, Chemometrics and Intelligent Laboratory Systems, 18 (1993) 251-263.



55. Branden, K. V. ve Hubert, M., Robustness Properties of a Robust PLS Regression Method. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.91.8504>, 1 Ekim 2012.
56. Alin, A. ve Masoom Ali, M., Improved Straightforward Implementation of A Statistically Inspired Modification of the Partial Least Squares Algorithm, Pakistan Journal Statistics , 28, 2 (2012) 217-229.
57. Avcı, E., Türkoğlu, İ. ve Poyraz, M., Darbeli Radar Sistemlerinde Bayes Karar Kuralı ile Hedef Sınıflama Uygulaması, Doğu Anadolu Bölgesi Araştırmaları, 2, 2 (2004) 25-29.
58. [http://www.cs.tut.fi/sgn/m2obsi/patrec/exercise\\_1.pdf](http://www.cs.tut.fi/sgn/m2obsi/patrec/exercise_1.pdf) Bayes Classifier. 1 Eylül 2012.

## **ÖZGEÇMİŞ**

Kübra EROĞLU, 1984 yılında Trabzon'da doğdu. Lise öğrenimini Trabzon Lisesi'nde tamamladı. 2003-2008 yılları arasında OnDokuz Mayıs Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Bölümü'nde Lisans Programı'nı tamamladı. 2010 yılında Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü'nde Yüksek Lisans Programı'na başladı. İyi seviyede İngilizce bilmektedir.