

**T.C.
TRAKYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**MAKİNE ÖĞRENMESİ YÖNTEMLERİNİN POLİSOMNOGRAFİK
VERİLERE UYGULANMASI**

GÜVEN ÇENTİK

**DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI**

**DANIŞMAN
YRD.DOÇ. DR. İLHAN UMUT**

EDİRNE, 2013

**MAKİNE ÖĞRENMESİ YÖNTEMLERİNİN POLİSOMNOGRAFİK
VERİLERE UYGULANMASI**

GÜVEN ÇENTİK

**DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ
ANA BİLİM DALI**

**2013
TRAKYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

T.Ü. Fen Bilimleri Enstitüsü Onayı

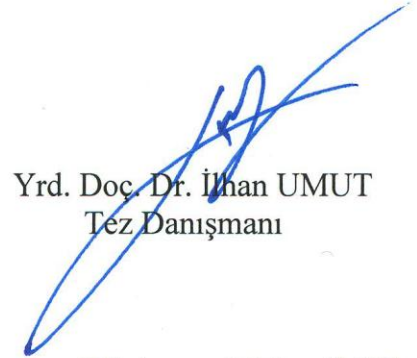
Prof. Dr. Mustafa ÖZCAN
Fen Bilimleri Enstitüsü Müdürü

Bu tezin Doktora tezi olarak gerekli şartları sağladığımı onaylarım.



Yrd.Doç.Dr. M.Tolga SAKALLI
Anabilim Dalı Başkanı

Bu tez tarafımızca okunmuş, kapsamı ve niteliği açısından bir Doktora tezi olarak kabul edilmiştir.



Yrd. Doç. Dr. İlhan UMUT
Tez Danışmanı

Bu tez, tarafımızca okunmuş, kapsam ve niteliği açısından Bilgisayar Mühendisliği Anabilim Dalında bir Doktora tezi olarak oy birliği ile kabul edilmiştir.

Jüri Üyeleri (Ünvan, Ad, Soyad):

Prof. Dr. A. Coşkun SÖNMEZ

Prof. Dr. Levent ÖZTÜRK

Doç. Dr. Yılmaz KILIÇASLAN

Doç. Dr. Erdem UÇAR

Yrd. Doç. Dr. İlhan UMUT

İmza



Tarih: 04 /10 / 2013

T.Ü.FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI DOKTORA PROGRAMI
DOĞRULUK BEYANI

İlgili tezin akademik ve etik kurallara uygun olarak yazıldığını ve kullanılan tüm literatür bilgilerinin kaynak gösterilerek ilgili tezde yer aldığını beyan ederim.

29/10/ 2013
Güven ÇENTİK

Doktora Tezi
Makine Öğrenmesi Yöntemlerinin Polisomnografik
Verilere Uygulanması
Trakya Üniversitesi Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

ÖZET

Uykuda Periyodik Hareket Bozukluğu (UPHB) Polisomnografi (PSG) kaydındaki bacak Elektromiyografisi (EMG) haricindeki diğer kanallar kullanılarak, dijital sinyal işleme yöntemleri ve makine öğrenmesi yöntemleri ile analiz edilmiştir. Bu hastalığın belirlenen özneliklere uygun makine öğrenmesi algoritmalarıyla sınıflandırılması amaçlanmıştır.

Çalışma UPHB hastalığı teşhisi konulmuş farklı yaş ve cinsiyetten 153 bireyin, PSG kayıtları kullanılarak gerçekleştirilmiştir. Hasta kayıtları (PSG) ortalama 7-8 saat sürmektedir. Çalışma ekibi tarafından dijital sinyal işleme ve makine öğrenmesi yöntemlerini kullanan bir yazılım geliştirilmiştir. Bu yazılım hasta kayıtlarını (PSG) 2, 5, 30 ve 60 saniyelik epok (bölüt) sürelerine parçalayarak analiz etmekte ve elde edilen sonuçları karşılaştırmaktadır. Her epok PSG kanallarının kendine ait farklı özneliklerini içermektedir. Bu öznelikler dijital sinyal işleme yöntemleri (Dalgacık ve Fourier) ve istatistiksel yöntemler kullanılarak elde edilip veritabanında saklanmıştır. Veritabanında saklanan öznelikler, farklı makine öğrenmesi algoritmaları kullanılarak UPHB sınıflandırılmıştır.

60 ve 30 saniyelik epoklar kullanılarak elde edilen sınıflandırma sonuçlarının birbirine çok yakın ve yüksek sınıflandırma oranlarına (% 79) sahip olduğu gözlemlenmiştir. 2 ve 5 saniyelik epoklar kullanılarak elde edilen sınıflandırma sonuçlarının da birbirine çok yakın olduğu ve düşük sınıflandırma oranlarına (% 63.93) sahip olduğu görülmüştür.

Makine öğrenmesi algoritmalarından 'lazy.Ibk' sınıflandırma algoritması diğer algoritmalara göre daha yüksek sınıflandırma oranı (% 79) ve düşük hata değerine (RMSE= 0.4652) sahiptir. Diğer yandan 'functions.Logistic' algoritması ise daha düşük sınıflandırma oranı (% 63.34) ve daha yüksek hata değerine (RMSE= 0.4741) sahiptir.

Yapılan analizler sonucunda, UPHB'nin bacak EMG'si kaydı kullanılmayarak, makine öğrenmesi algoritmalarıyla, % 79 gibi yüksek bir oranla sınıflandırılabilceği tespit edilmiştir.

Yıl: 2013

Sayfa Sayısı: 89

Anahtar Kelimeler: Uykuda Periyodik Hareket Bozukluğu (UPHB), Sayısal İşaret İşleme, Makine Öğrenmesi, Polisomnografi (PSG)

Doctorate Dissertation
Application of Machine Learning Methods to
Polysomnography Datas
T.U. Institute of Natural and Applied Sciences
Department of Computer Engineering

ABSTRACT

The Periodic Leg Movement Disorder (PLMD) has been analyzed with the DSP (Digital Signal Processing) and machine learning methods by means of other channels apart from the leg electromyography in records polysomnography (PSG). This study aimed to classify PLMD by means of the selected attributes in accordance with the machine learning algorithms.

This study was carried out by means of PSG recordings among 153 people, defined by age and sex. PSG recordings take around 7-8 hours. A software program that uses the Digital Signal Processing (DSP) and machine learning techniques was developed by the working team. This software analyzes PSG recordings in 2, 5, 30, and 60-second epochs and compares the obtained results. Each epoch contains different attributes specific to its own. These attributes are deposited in the database with statistical and DSP analysis. The obtained attributes in the database are used to classify PLMD by using different machine learning algorithms.

It is observed that classification results that obtained by using 60 and 30-second epochs are very close to each other and have high classification rates (79 %). It is also observed that classification results that obtained by using 2 and 5-second epochs are very close to each other but have low classification rates (63.93 %).

'lazy.Ibk' classification algorithm in machine learning algorithms has higher classification rate (79 %) and lower error value (RMSE= 0.4652) than other algorithms. On the other hand 'functions.Logistic' classification algorithm has lower classification rate (63.34 %) and higher error value (RMSE= 0.4741).

As a result of the analyzes, PLMD using PSG record other than the leg EMG channels, classified as high as rate (79 %) by the machine learning algorithms.

Year: 2013

Number of Pages: 89

Keywords: Periodic Leg Movement Disorder (PLMD), Digital Signal Processing (DSP), Machine Learning, Polysomnography (PSG)

TEŐEKKÜR

Doktora alıőmalarım sırasında yardımlarını esirgemeyen, akademik bilgi ve tecrübelerini örnek aldığım hocalarım, Sayın Yrd. Do. Dr. İlhan Umut'a, Sayın Do. Dr. Erdem Uar'a, Sayın Prof. Dr. Levent Öztürk'e, sonsuz teőekkürlerimi saygıyla sunarım.

alıőmalarım sırasında istatistik konusunda benden yardım ve desteklerini esirgemeyen Sayın Do. Dr. Necdet Süt'e teőekkürlerimi saygıyla sunarım.

alıőmalarımda PSG verilerinin temini konusunda bana yardımcı olan Sayın Necati Aslan'a teőekkürlerimi sunarım.

Ayrıca, doktora alıőmalarım boyunca bana destek veren, sabır gösteren her zaman yanımda olan saygı deđer annem ve babama teőekkürü bir bor bilirim.

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER.....	iv
SİMGELER VE KISALTMALAR.....	vi
ŞEKİLLER LİSTESİ.....	viii
TABLolar LİSTESİ.....	x
1 GİRİŞ	1
1.1 Konu	1
1.2 Amaç.....	1
1.3 Önem	2
1.4 Kapsam	2
1.5 Literatür Özeti.....	3
1.5.1 Sinyal İşleme Teknikleri ile İlgili Literatür Atıfları	3
1.5.2 Makine Öğrenmesi ile İlgili Literatür Atıfları	4
2 MATERYAL METOD	8
2.1 Uyku	8
2.1.1 Uyku Evreleri	8
2.2 Polisomnografi (PSG)	9
2.2.1 Polisomnografi Nasıl Yapılır?	9
2.3 UPHB Tanı Kriterleri	18
2.4 Veritabanı	18
2.5 Çalışmada Kullanılan Dijital Sinyal İşleme Metotları	19
2.5.1 Fourier Analiz	19
2.5.2 Dalgacık Dönüşümü	23
2.6 Makine Öğrenmesi Nedir?	27
2.6.1 Makine Öğrenmesinin Amaçları	29
2.7 Makine Öğrenmesi Algoritmaları	31
2.7.1 En Yakın K-Komşu Algoritması	31
2.7.2 Naive Bayesian Yöntemi.....	33
2.7.3 Yapay Sinir Ağları	34

2.7.4	SOM Yapay Sinir Ağları: Özdüzenleyici Haritalar.....	38
3	GELİŞTİRİLEN ANALİZ YAZILIM MODÜLLERİ.....	42
3.1	Verilerin Elde Edilmesi	42
3.2	Veritabanı Modülü.....	44
3.2.1	Yeni Veritabanı Oluşturma.....	47
3.2.2	Analiz Tablosu Oluşturma	49
3.3	Dijital Sinyal İşleme Modülü	50
3.3.1	Öznitelikler	55
3.4	Artefakt Modülü	57
3.5	XML Modülü.....	59
3.6	Makine Öğrenmesi Modülü	62
3.6.1	WEKA	66
4	PSG VERİLERİ ÜZERİNE UYGULAMA.....	68
4.1	Modelin Kurulması	69
4.1.1	Sınıflandırma Başarımı Ölçüm Metrikleri.....	69
4.2	Modelin Değerlendirilmesi	70
4.2.1	Karışıklık Matrisi.....	71
4.2.2	İşlem Karakteristiği Eğrisi (ROC-AUC)	72
4.3	Uygulama Sonuçları.....	76
4.4	Uygulama Sonuçlarının Değerlendirilmesi.....	77
4.5	Öneriler	78
	KAYNAKLAR.....	80

SİMGELER VE KISALTMALAR

AFD	: Ayrık Fourier Dönüşümü
ANFIS	: Adaptive Neuro-Fuzzy Inference System
AUC	: Area Under Curve
CM	: Karışıklık Matrisi
DD	: Dalgacık (Wavelet) Dönüşümü
DLL	: Dynamic-Link Library
DSİ	: Dijital Sinyal İşleme
EDF	: European Data Format
EDF+	: European Data Format Plus
EEG	: Elektroensefalografi
EKG	: Elektrokardiyografi
EMG	: Elektromyografi
EOG	: Elektrookülografi
GAM	: Generalized Additive Model
GLM	: Generalized Linear Model
GML	: Generalize Markup Language
HBS	: Huzursuz Bacak Sendromu
HFD	: Fast Fourier Transform (FFT)
IBM	: International Business Machines
KNN	: K-Nearest Neighbors
KSFD	: Kısa Süreli Fourier Dönüşümü
LDA	: Linear Discriminant Analysis
LDA	: Linear Discriminant Analysis
LOC	: Left Outer Canthus
MGG	: Mixture of Generalized Gaussian
MOG	: Mixture of Gaussian
Non-REM	: Non Rapid Eye Movement
ÖDH	: Özdüzenleyici Harita
PCA	: Patient Controlled Analgesia

PLM	: Periodic Leg Movement
PSA	: Prostat Spesifik Antijen
PSG	: Polisomnografi
QDA	: Quadratic Diskriminant Analysis
RBF	: Radyal Tabanlı Fonksiyon
REM	: Rapid Eye Movement
RERA	: Respiratory Effort Related Arousal
ROC	: Right Outer Canthus
RPART	: Recursive Partitioning and Regression Tree
SGML	: Standard Generalized Markup Language
SOM	: Self-Organizing Maps
STFT	: Short-Time Fourier Transform
SVM	: Support Vector Machine
UPHB	: Uykuda Periyodik Hareket Bozukluğu
W3C	: World Wide Web Consortium
XML	: Extensible Markup Language
YSA	: Yapay Sinir Ağları

ŞEKİLLER LİSTESİ

Şekil 2.110-20 sistemine göre elektrot yerleşim düzeni [10].....	10
Şekil 2.2 Farklı bantlardaki EEG dalgaları [40]	12
Şekil 2.3 Örnek bir EMG sinyali	13
Şekil 2.4 EKG işaretinin bölümleri [44]	14
Şekil 2.5 Termokupl ve termistörler ile hava akımı ölçme cihazları [46]	14
Şekil 2.6 Nazal basınç transdüseri ile hava akımı ölçülmesi [46].....	15
Şekil 2.7 Oksihemoglobin saturasyonu ölçme metodları [41].....	16
Şekil 2.8 Pozisyon sensörü [41].....	16
Şekil 2.9 Horlama ölçümünde mikrafon kullanılması [41].....	17
Şekil 2.10 Zaman-Frekans-Genlik düzlemi [54].....	19
Şekil 2.11 KSFD ile Dalgacık Dönüşümü arasındaki farklılıklar [60]	24
Şekil 2.12 Dalgacık (Daubechies) [60]	25
Şekil 2.13 Haar dalgacığının grafiksel gösterimi.....	26
Şekil 2.14 Dalgacık Dönüşümü'nün seviyesi artırıldıkça, yaklaşım katsayıları ve detaylı katsayıların nasıl üretildiğini gösteren blok diyagram	27
Şekil 2.15 KNN sınıflandırıcı örneği [84].....	32
Şekil 2.16 Yapay nöron [98]	34
Şekil 2.17 Yapay sinir ağı modeli [98]	35
Şekil 2.18 Toplam ve aktivasyon fonksiyonlu bir yapay sinir ağı [98]	36
Şekil 2.19 Çok katmanlı ileri ve geri beslemeli ağların yapıları [101]	37
Şekil 2.20 N adet özneteliğe sahip girdi veri seti için tasarlanmış 3x3 boyutlu SOM'un yapısı [105]	39
Şekil 2.21 7x7 SOM üzerinde komşuluk ilişkilerinin dikdörtgensel topoloji ve 2, 1, 0 yarıçapları için gösterimi [105].....	40
Şekil 3.1 PSG kaydı örneği	43
Şekil 3.2 Veritabanı Modülünün ekran görüntüsü	46
Şekil 3.3 Yeni veritabanı oluşturma akış şeması	48
Şekil 3.4 Epok süresi 5 sn olan örnek bir analiz ve veritabanında bulunan, tablolar ve arasındaki ilişkiler	49

Şekil 3.5 DSP Modülünün ekran görüntüsü.....	52
Şekil 3.6 DSP Modülünün akış şeması	54
Şekil 3.7.Göz kırpması sonucu EEG dalgasında oluşan artefakt [111]	58
Şekil 3.8 Artefakt Modülünün arayüzü	58
Şekil 3.9 Örnek bir XML dokümanı	60
Şekil 3.10 XML Modülünün arayüzü	61
Şekil 3.11 XML (Skorlama) Modülü arayüzü	62
Şekil 3.12 Makine Öğrenmesi Modülünün oluşturduğu örnek dosyalar	63
Şekil 3.13 Makine Öğrenmesi Modülünün ekran görüntüsü	65
Şekil 3.14 WEKA’da Applications Menüsü	66
Şekil 3.15 WEKA’da veri seçimi.....	67
Şekil 4.1 ‘Arff’ dosya formatındaki niteliklerin bazıları ve niteliklerin sayısı	68
Şekil 4.2 ‘rules.JRip’ algoritmasının sonuçları	71
Şekil 4.3 İdeal (A) ve kötü performans (B) gösteren iki testin ROC eğrisi.....	74

TABLolar LİSTESİ

Tablo 2.1 Standart polisomnografi parametreleri	10
Tablo 2.2 EEG işaretlerinin frekans bantları [38].....	11
Tablo 3.1 Farklı epok sürelerine göre PLM'nin olduğu ve olmadığı, artefaktlı ve artefaktsız epoksayıları.....	44
Tablo 3.2 Modülde kullanılan öznitelikler ve değerleri (kesirli sayı:1, tamsayı: 2).....	56
Tablo 4.1 60 sn'lik epok süreleri ile makine öğrenmesi algoritmalarının performans analizi	75
Tablo 4.2 5 sn'lik epok süreleri ile makine öğrenmesi algoritmalarının performans analizi	75

1 GİRİŞ

1.1 Konu

Polisomnografi (PSG) uyku bozuklukları ile ilgi hastalıkların teşhisinde kullanılan bilgisayarlı kayıt sistemidir. Uyku esnasında hastadan farklı kanallardan biyoelektrik sinyaller ölçülerek kayıt edilmektedir. Biyoelektrik sinyaller, bilgisayar ortamında hekim tarafından skorlanarak uyku hastalıklarının teşhisinde kullanılmaktadır. Bu hastalıklardan birisi de uykuda periyodik bacak hareketleridir. Uykuda periyodik bacak hareketleri uyku sırasında kısa kas kasılmalarıdır. Kasılmalar her 30 saniyede bir, sıklıkla bir saat veya daha uzun bir süre boyunca devam eder. UPHB bacak EMG'si kullanılarak tespit edilmektedir. Bu çalışmada UPHB bulunan 153 hastanın PSG kaydı bilgisayar ortamında analiz edilmiştir. Uykuda periyodik hareket bozukluğunun (UPHB) bacak EMG'si dışındaki diğer PSG sinyalleri ile ilişkisi makine öğrenmesi teknikleri kullanılarak incelenmiştir.

1.2 Amaç

Bu çalışmanın amacı UPHB'nin, bacak EMG'si haricinde PSG kayıtlarında bulunan diğer kanallarla ilişkili olup olmadığını tespit etmektir. Çalışma ekibi tarafından bu amaçla, UPHB teşhisi konulmuş bireylerin PSG kayıtlarını analiz etmek için dijital sinyal işleme, makine öğrenmesi ve istatistiksel yöntemleri kullanan modüler bir bilgisayar yazılımı geliştirilmiştir.

1.3 Önem

Birçok araştırma uyku yetersizliğinin insan hayatında önemli bir sorun olduğunu göstermektedir. Uyku yetersizliği olan kişiler, sürüş simülatörü, göz-el performans koordinasyon görevi ile test edilmiş ve alkollü kişilerden daha kötü performans sergiledikleri görülmüştür [1].

Uykuda Periyodik Hareket Bozukluğu [Periodic Limb Movement Disorder (PLMD)] gündüz aşırı uykusuzluk, geceleyin uyanma, uyku döngüsünde bozulma, uykudan sık uyanma, yatak eşini tekmeleme gibi olumsuz etkileri olan bir hastalıktır. Hastaların bazıları kol ve bacak atmaları ile uyanmadan yakınmakta, ayrıca çoğunda sabahları bacaklarda ağrı, dinlenmemiş ve yorgun uyanma veya nadiren gündüz uykululuk hali görülmektedir [2].

UPHB, kol ve bacaklarda tekrar eden kasılma hareketlerine neden olur. 20 saniye ile 40 saniye arasında değişen bu hareketler uyku kalitesini düşürmektedir. Yapılan bir çalışma, UPHB ve Huzursuz Bacak Sendromu (HBS)'nin 60 yaş üzerindeki hastalarda uykusuzluğa neden olduğunu göstermiştir [1].

1.4 Kapsam

UPHB 30 yaş öncesinde nadir olarak gözlenmektedir. Sıklığı ve şiddeti yaşla doğru orantılı olarak artan bir hastalıktır. UPHB'nin varlığını değerlendiren en geniş araştırma, 18.980 kişi üzerinde yapılmıştır. Araştırma yaşları 15–100 arasında değişen genel nüfusta yapılmış ve prevalansın % 3,9 olduğu sonucuna varılmıştır. Araştırmada hastanın cinsiyeti, yaşı, kafein tüketim miktarı, stres yoğunluğu ve mental hastalıkları olup olmadığı gibi birçok kriter baz alınarak bunların UPHB ile ilişkisi değerlendirilmiştir. Prevalans değerinin 55 yaş üstü bireyler için % 29, 65 yaş üstü bireyler için ise % 44 olduğu sonucuna ulaşılmıştır. UPHB, her iki cinsiyette eşit sıklıkla görülmektedir. Nadiren ailesel özellik göstermektedir [3].

UPHB'nin skorlaması ve Polisomnografi (PSG) deęerlendirmesi 1993'te standart hale getirilmiřtir [4]. Bu ölçütler hala UPHB arařtırmalarında altın standart olarak kabul edilir. Yalnız, UPHB'nin skorlama ölçütü sınırlı sayıda gözlemlere baęımlıdır ve farklı hasta grupları arasında UPHB davranıřları henüz geçerlilik kazanmamıřtır. UPHB'nin oluřma frekansı, yani UPHB indeksini belirlemek için toplam uyku zamanının saat başına düşen UPHB sayısı hesaplanır. UPHB PSG kayıtlarına ait bacak Elektromiyografi (EMG) iřaretlerinin uzman hekimler tarafından skorlanması ile teřhis edilir [5].

1.5 Literatür Özeti

1.5.1 Sinyal İşleme Teknikleri ile İlgili Literatür Atıfları

1946 yılında Dennis Gabor, iřaretleri küçük parçalara ayırarak incelemiřtir. Fourier dönüşümünün eksik kaldığı noktaları gidermeye çalışmıřtır. Gaborun kullandığı bu metoda, Kısa-Zaman Fourier Dönüşümü (FD) denilmektedir. Kısa-Zaman Fourier Dönüşümü, iřareti zaman ve frekans domeninde incelemektedir [6].

Grap, 1995 yılında yaptıęı çalışmada, Fourier analizlerinin ve Dalgacık teorisinin gelişimini incelemiřtir. Sinyal işleme konusunda birçok karşılařtırma yapmıř, yeni gelişmekte olan dalgacık analizinin uygulama alanlarını ortaya koymuřtur [7].

1997 yılında Dograhmi, EEG sinyallerinin spektral analiz temelli sınıflandırılması konusunda çalışmıřtır. EEG elektrot sayısının ve uygulanan spektral analiz yöntemi özelliklerinin bir sınıflandırıcı gibi kullanılmasını incelemiřtir [8].

Jung ve arkadaşlarının 1997 yılında yaptıkları çalışma, uyanıklık ve uyku hallerini YSA (Yapay Sinir Ağları) ile sınıflandırma konusundadır. YSA kullanarak sınıflandırmada, ham EEG sinyallerini kullanmak yerine FD'nü kullanmıřlardır [8].

1998 yılında Miner, Fourier tekniklerinin ve Dalgacık teorisinin analizlerinin gelişimini incelemiřtir. İřaret işleme konusunda çeřitli karşılařtırmalar yapmıřtır. Çalışmasında Dalgacık teorisi konusunda, ayırık ve sürekli dalgacık analizi algoritmalarını sunmuřtur [9].

2004 yılında Kıymık ve arkadaşları, EEG işaretleri üzerinde epilepsi hastalığını Kısa Zaman Fourier Dönüşümü ve Dalgacık Dönüşümünü eş zamanlı bir uygulama yaparak incelemiştir. Dalgacık dönüşümünün Kısa Zaman Fourier Dönüşümüne kıyasla daha iyi sonuçlar verdiğini göstermişlerdir [10].

2005 yılında Hatice Batar yüksek lisans tezinde, EEG sinyali gibi durağan olmayan rastgele seçtiği sinyalleri uyuklama, uyanıklık ve uyku durumlarında gösterdikleri farklılıkları YSA kullanarak sınıflandırmıştır [10].

2007 yılında Atılay Altınbaş yüksek lisans tezinde, hastanede pratik çalışmalar sonucunda elde edilen kas hareketleri ile oluşan elektriksel aktivitenin algılanmasını, Kısa Zaman Fourier Dönüşümü ve DD tekniklerini kullanarak belli katsayılarla ortaya koymuştur. Kısa Zaman Fourier Dönüşümü ve DD kullanarak çeşitli kriterlere göre sonuçlar elde etmiştir [11].

2008 yılında Ömer Demir yüksek lisans tezinde, farklı derecelerde kafa travması geçirmiş 4'ü kadın, 6'sı erkek toplam 10 hastanın EEG kayıtlarını travmayı takip eden ilk 24 saat içinde kayıt etmiştir. MatLab programını kullanarak ADD metodu ile EEG kayıtlarına ait periyodikliğini yakalayarak ve EEG sinyallerinin frekansını hesaplayarak travma derecesi belirleyecek bir metod ortaya koymuştur [6].

1.5.2 Makine Öğrenmesi ile İlgili Literatür Atıfları

2002 yılında Chan ve arkadaşları, görme alanı testi için kullanılan standart otomatik perimetri verileri üzerinden STATPAC endeksleri ile makine öğrenmesi yöntemlerinden MOG, MGG, LDA ve QDA, Parzen Window gibi yöntemleri kullanarak karşılaştırma yapmışlardır. Makine öğrenmesi tekniklerinin klasik STATPAC teşhis tekniklerine göre üstün olduğu sonucunu ortaya koymuşlardır [12].

2003 yılında Selmin Danış tarafından hazırlanan yüksek lisans tezinde, 50 yaş ve üstündeki erkeklerde prostat kanseri teşhisi için Yapay Sinir Ağları, Karar Ağacı ve K-En Yakın Komşu Algoritmalarını, veri madenciliği algoritmalarını kullanmıştır. Tezinde 139 hastanın verilerini kullanarak bu üç veri madenciliği algoritmalarını birbiriyle kıyaslamıştır [13].

2004 yılında Muhammed Şakir Toprak tarafından hazırlanan yüksek lisans tezinde, tıp alanında cerrahi müdahalelerde kalp krizi riskinin tespit edilmesi konusunda makine öğrenmesi teknikleri ile hekimlerin karar vermesine yardımcı olacak bir sistem

tasarlanmıştır. Genetik algoritma ve yapay sinir ağı yöntemlerini kullanarak hekimler olmadan kendi başına öğrenen ve karar veren bir sistem geliştirmiştir [14].

2005 yılında Yusuf Uzun tarafından hazırlanan yüksek lisans tezinde, ameliyat sonrasında hastalardan aldığı tıbbi veri kümelerinden 9 özellik baz alınarak Jrip, Ridor, SMO, J48, OneR, Karar Ağaçları, Naive Bayes, LMT, NNge, KStar, IBk, PART Conjunctive Rule, Decision Tables, gibi makine öğrenmesi algoritmaları ve bulanık mantık ile sınıflandırma analizleri ve kural öğrenimi yaparak algoritmaları ve bulanık mantık sonuçları performans değerlendirmesine tabi tutulmuştur [15].

2005 yılında Eliash ve arkadaşları, makine öğrenme yöntemlerini birbiri ile karşılaştıran çalışmalarında göz tomografisi ölçümlerinden elde edilen sonuçlar ile sınıflandırma yapmışlardır. Kullanılan makine öğrenmesi yöntemleri; SVM, LDA, RPART, GLM ve GAM'dır. Çalışmalarında 47 glokom hastası ve 42 sağlıklı göz içerisinde en yüksek değeri SVM modeli vermiştir. Ayrıca çalışmalarında 47 hastanın 27'si erken glokom, 20'si ileri glokom olgulu olarak aralarında yapılan sınıflamada ise en yüksek başarıyı GAM'ın verdiğini ortaya koymuşlardır [16].

2007 yılında İnci Zaim Gökbay yüksek lisans tezinde, makine öğrenmesi algoritmalarıyla meme kanserinin erken teşhisinin gerçekleştirilebilmesini araştırmıştır. Tezi kapsamında geliştirdiği sistemi mamografi görüntülerine uygulamış ve önerilen metodun, radyologların mamografi görüntülemeye kullanabileceği alternatif bir yöntem olduğu sonucuna varmıştır [17].

2007 yılında Bock ve arkadaşları tarafından, Renkli fundus resimleri ile Almanya'da yapılmış bir çalışmada obje ve yüz tanımadaki gibi doğrudan fundus resimleri ile sınıflandırma yöntemine gidilmiş, sınıflandırma için resimlerle ön çalışma yapıldıktan sonra Naive Bayes, K-Nearest Neighbor ve Support Vector Machine kullanılarak karşılaştırılmıştır. 50 hasta ve 50 sağlıklı bireye ait fundus resimlerinde yapılan sınıflandırmada SVM yöntemi en başarılı performansı vermiştir [18].

2008 yılında M. Fatih Amasyalı tarafından hazırlanan doktora tezinde, ilaç tasarımı sürecinin ve maliyetinin önemli bileşenlerinden olan ilaç moleküllerinin seçilmesi işlemi, sınıflandırma, kümeleme, özellik seçimi/çıkarması, eğri uydurma gibi problemlerden birkaçını makine öğrenmesi metotları yardımıyla ilaç tasarımı sürecini ve maliyetini azaltabilecek bir çalışma ortaya koymuştur. Mevcut

algoritmalarla, çeşitli kümeleme performans kriterlerine göre yarışan sonuçlar elde etmiştir [19].

2009 yılında Kahraman doktora tezinde, geleneksel öğretim sistemlerinde ortaya çıkan problemlerin çözümüne yönelik, Bayes teoremini kullanarak bir öğretim aracı geliştirmiştir. Geliştirilen araç sayesinde öğrencilerin geleneksel sistemlere kıyasla daha başarılı olduğu sonucuna varmıştır [20].

2010 yılında Çağdaş Çalış yüksek lisans tezinde, ANFIS ve Rough Set veri madenciliği yöntemlerini kullanmıştır. Cinsiyet, boy kilo endeksi, hba, tansiyon, şeker, kolestrol, kandaki keton, kandaki protein, kandaki mikalb, kandaki mikros, genetik ve ilaç dozunu giriş parametreleri olarak kullanmıştır. Dozaj planlamasını Coversyl, Monopril, Tenoretic ve Atacand ilaçları üzerinde yapılmıştır. ANFIS'in RSES'den daha iyi sonuçlar verdiği gözlenmiştir. ANFIS'in, dozaj planlaması yapılırken en güvenilir yöntem olduğu sonucuna ulaşmıştır [21].

2011 yılında Aslı Uyar Özkaya doktora tezinde, tüp bebek tedavisinde klinik başarı oranlarının artırılması için karar verme problemleri, yapay öğrenme bakış açısı ile ele alınmıştır. Bayes Ağları yöntemi kullanılarak embriyo gelişim süreci modellenmiştir. Özkaya, koşullu olasılık tablolarındaki parametrelerin daha iyi öğrenilebilmesi için yeni bir yöntem önermiştir. Çalışmasında, önerilen yöntemler kullanıldığında tahmin performansının arttırılabileceğini ortaya koymuştur. Klinik açıdan çoğul gebeliklerin azaltılması, embriyo kayıplarının azaltılması ve transfer iptallerinin engellenmesini sağlamaya çalışmıştır [22].

2011 yılında Mustafa Karabulut doktora tezinde, üç farklı hesaplama yöntemi önermiştir. Önerilen birinci yöntemde, verilen tüm w -mer'leri kümeleme ve z -score testini kullanarak istatistiksel olarak ilginç bir yerel hizalama bulmaya çalışmıştır. Self-organizing map, Fuzzy C-Means, Kmeans ve Expectation Maximization algoritmalarını kullanmıştır. İkinci yöntem ise birinci yönteme oldukça benzemekle beraber, farklı olarak, pozisyon ağırlık matrislerinden oluşan yerel hizalamaları iyileştirmek için Bayes teoremi temelli kümeleme sonrası optimizasyon prosedürü kullanmıştır. Geliştirdiği metotlar, ilgili literatüre yeni katkılar sağlamak amacıyla düşük ve yüksek canlı DNA'sı içeren pek çok veri seti kullanılarak değerlendirilmiştir. Literatürden MEME ve MDSscan gibi gelişmiş metotlarda kıyaslama yapmıştır. Deneysel sonuçlar önerilen

metotların DNA motif-bulma işi için oldukça umut vaat edici olduğunu göstermiştir [23].

2012 yılında Şerife Hacıfendioğlu tarafından hazırlanan yüksek lisans tezinde, görme kayıpları oluşmadan glokom hastalığının erken teşhisine yardımcı olacak şekilde makine öğrenmesi yöntemleri ile sınıflandırma çalışmaları yöntemlerini birbirleri ile karşılaştırmıştır. Kullandığı yöntemlerin performansları X-Validation ile belirlenmiş ve en yüksek sınıflandırma başarısının Destek Vektör Makineleri ile elde edileceğini göstermiştir [24].

2012 yılında Ümit Can Kundereli yüksek lisans tezinde, Elektroansefalografi (EEG) cihazlarından almış olduğu verileri, geliştirdiği yazılım sayesinde analiz etmiş, ayırık fourier ve dalgacık dönüşümüne tabi tutmuştur. Diğer yapılan çalışmalardan farklı olarak 19 kanallı 30 dakikalık rutin çekimler sonucu elde edilen veriler kullanılmış olup bu veri setine uygun öznelik parametreleri seçilmiştir. Veriler üzerinde çeşitli sınıflandırma algoritmaları uygulanarak; yüksek doğruluk oranı ile epileptik aktiviteyi teşhis eden algoritmalar belirlenmiştir. Böylece nöroloji uzmanlarının ileride epileptik aktiviteyi belirlemede teşhis süresini kısaltacak ve teşhisin doğruluğunu artıracak bir model oluşturulmaya çalışılmıştır [25].

2012 yılında Alper Vahaplar doktora tezinde, veri madenciliği yöntemlerinin bazılarını dikotik dinleme testi sırasında kaydedilen EEG verileri üzerinde uygulamayı hedeflemiştir. EEG verisini detaylı olarak incelemiş, analiz edip parçalara ayırmış ve etiketlendirmiştir. Farklı uyaranların etkisiyle oluşan tepkileri ve farklı elektrotlardaki sinyalleri karşılaştırmak ve benzerlik ya da benzemezliği tespit etmek üzere ZM istatistiğini temel araç olarak kullanmıştır [26].

2 MATERYAL METOD

2.1 Uyku

Uyku, davranışsal olarak normal, tekrarlayan ve dış çevrenin algılanması ve cevap verilmesi yetisinin geçici süreyle kaybı olarak tanımlanabilir. İstemli kas hareketleri büyük oranda azalır ve sakin bir duruşa geçilir. Beyin tarafından üretilir ancak tüm vücutta fizyolojik etkileri mevcuttur. Uyanıklığın olmadığı pasif bir durum inanın aksine uyku nöral fonksiyonlar üzerinde aktif ve dinamik değişikliklere sebep olmaktadır [27].

Uyku evrelerinin standart terminolojisi hakkında 1968'de Rechtschaffen ve Kales (R&K) editörlüğünde, teknik özellikler ve skorlamaya yönelik bir kitapçık hazırlanmıştır [28]. Yakın döneme kadar uyku çalışmalarının skorlanması ve değerlendirilmesi bu kılavuza göre yapılırken 2007 yılında çıkan Amerikan Uyku Tıbbi Cemiyeti (AASM) kılavuzu, uyku evreleri de dahil olmak üzere birçok değişiklik teklif etmiştir [29].

2.1.1 Uyku Evreleri

Kişi uyku esnasında bir dizi evreden geçmektedir. Bu evrelerin sayısı ve sırası uyku kalitesini değerlendirmek açısından önemlidir. Uyku temel olarak 2 evreden oluşur. Bu evreler Non-REM ve REM evreleridir. Non-REM evresi de 3 evreden oluşur [30]. REM evresi EOG sinyalindeki değişiklikler ile tespit edilmektedir. REM evresinde çok hızlı göz hareketleri belirmektedir. Non-REM evresi kendi içinde dört evreye ayrılır. Bu evreler Non-REM1, Non-REM2, Non-REM3 ve Wake (W) evreleridir [31].

2.2 Polisomnografi (PSG)

PSG, uyku laboratuvarlarında uygulanan ve uykuda solunumsal problem yaşayan hastalarda tanı koymada kullanılan standart testtir. Gece boyunca uyku esnasında vücudun ürettiği biyoelektrik sinyaller devamlı, gerçek zamanlı olarak kayıt edilir [32]. PSG, günümüzde uyku laboratuvarlarında kullanılan bir test yöntemidir [30].

2.2.1 Polisomnografi Nasıl Yapılır?

PSG 6-8 saat arasında tüm gece süresince yapılmaktadır. Uyku odalarında kamera ve ses kayıt cihazları da bulunmaktadır. PSG uyku teknisyeni tarafından gerçekleştirilmektedir. Hastaya biyoelektrik sinyal elektrotları uyku teknisyeni tarafından yerleştirilmektedir. Bu işlem 30-60 dakika süre alır. Teknisyen uyku odasındaki hastayı odadaki kamera yardımıyla görmekte mikrofonlar yardımıyla duyabilmektedir. PSG verileri günümüzde bilgisayara kayıt edilmektedir [33]. Gece vardiyasında çalışanlar için test gündüz yapılacaksa, testin süresi ve protokolü gece uygulamasındaki gibi olmalıdır [34].

Polisomnografide 6 kanal EEG bazen 2 kanal EEG, 1 kanal EMG, 2 kanal EOG, genelde rutin olarak 12-16 kanallı olarak kayıt edilmektedir. Elde edilen sinyaller düşük voltajlı olduklarından amplifikasyon (yükseltme) işlemine tabi tutulurlar ve daha net bir kayıt elde etmek için bazen filtreler kullanılır [27].

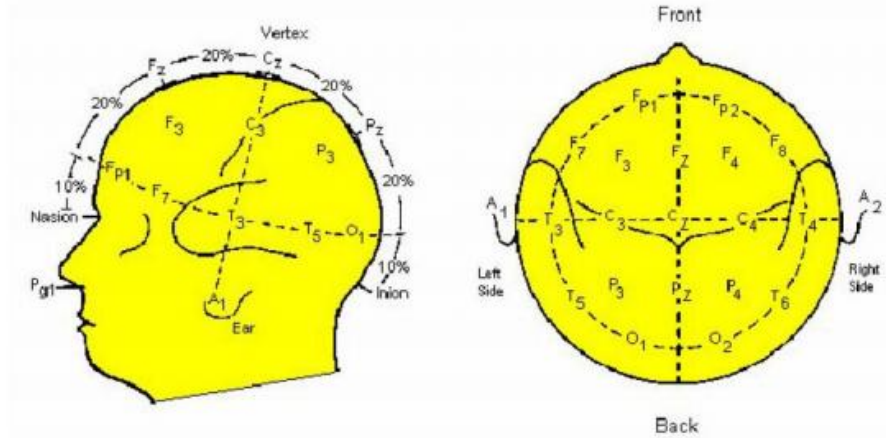
Standart polisomnografi sinyalleri Tablo 2.1’de gösterilmiştir [35].

Tablo 2.1 Standart polisomnografi parametreleri

Standart Polisomnografi Sinyalleri
1. Elektroensefalografi (EEG)
2. Elektrookulografi (EOG)
3. Elektromyografi (EMG-submentalis)
4. Ora-nasal hava akımı
5. Torako-abdominal hareketler
6. Oksijen satürasyonu
7. Elektrokardiyografi (EKG)
8. Elektromyografi (EMG-tibialis)
9. Vücut pozisyonu

2.2.1.1 Elektroensefalografi (EEG)

Beynin faaliyetleri sırasında elektriksel potansiyeller oluşur. Bu potansiyeller EEG ile kaydedilir. Hastanın başına uluslararası 10-20 sistemine göre elektrotlar yerleştirilerek EEG'deki değişiklikler kayıt edilir. Aşağıda 10-20 sistemi Şekil 2.1'de gösterilmiştir [10].



Şekil 2.110-20 sistemine göre elektrot yerleşim düzeni [10]

EEG sinyalinin içerdiği frekans aralığı oldukça geniş olmasına rağmen klinik ve fizyolojik ilgi 0,5 ile 30 Hz arasına yoğunlaşmıştır. EEG sinyalinin genliği, beyin aktivitesine bağlı olarak 1 mikrovolt ile 400 mikrovolt arasında değişir [36].

2.2.1.1.1 EEG frekansları

EEG, EKG ve EMG sinyalleri kapsadığı frekanslara göre sınıflandırılmaktadır. EEG sinyallerinin frekansı beyin aktivitesi ile değişiklik göstermektedir [37].

EEG işaretleri, kapsadıkları frekans bantlarına göre isimlendirilirler. Aşağıda Tablo 2.2’de EEG sinyallerinin frekans bant aralıkları gösterilmiştir.

Tablo 2.2 EEG işaretlerinin frekans bantları [38].

EEG Frekans Bantları	Frekans Bantları(Hz)	Genlik(mikrovolt)
Delta	0.5-4	20-400
Teta	4-8	5-100
Alfa	8-13	2-10
Beta	13-30	1-5
Gama	30<	1<

Delta Dalgaları: Frekansı 4 Hz’den küçük olan beyin dalgalarıdır. Genlikleri 20-400 mikrovolt arasındadır. Delta dalgalar derin uykuda, zaman zaman uyanık haldeyken ve genel anestezi gibi beynin düşük aktivite gösterdiği anlarda görülmektedir [39].

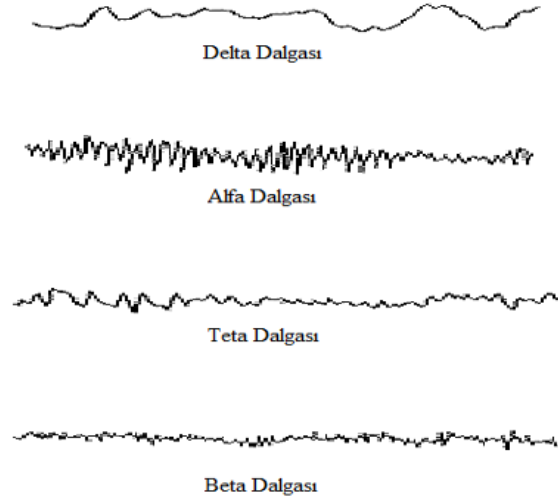
Teta Dalgaları: Frekansı 4–8 Hz arasındadır. Genlikleri 5 ila 100 mikrovolt arasında değişmektedir. Normal bireylerde; rüyalı uyku, orta derinlikte anestezi, uykunun başlangıç safhalarında, derin düşünme, duygusal gerginlik ve düş kırıklığı gibi beynin düşük aktivite gösterdiği zamanlarda görülmektedir [39].

Alfa Dalgaları: Frekansı 8 ila 13 Hz arasındaki beyin dalgalarıdır. Genlikleri 2 ila 10 mikrovolt arasında değişiklik gösterir. Başın arka kısmından elde edilmektedir. Genellikle beynin oksipital bölgesi üzerinde bulunur. Alfa Dalgaları sıklıkla sinüs dalgası şeklindedir. Nadiren keskin köşeli şekillerde gözlenmektedir [39].

Beta Dalgaları: Frekansları 13 ila 26 Hz aralığındaki beyin dalgalarıdır. Genlikleri 1 ila 5 mikrovolt arasında değişir. Çok yüksek beta seviyesi, panik halinde ortaya çıkar. Ritmik beta durumu, çoğunlukla beynin ön ve merkez bölgelerinden elde edilir [39].

Gamma Dalgaları: Frekansları 30 Hz’in üstündeki beyin dalgalarıdır. Genlikleri, 2 mikrovolttan küçüktür. Kafanın merkezinde, genlikleri daha büyüktür. Uykunun karakteristik belirtisini taşırlar ve beynin her tarafında bulunan tek frekans band grubudur [39].

Aşağıda Şekil 2.2’de EEG işaretlerine ait farklı frekans aralıklarındaki dalgalar gösterilmiştir.



Şekil 2.2 Farklı bantlardaki EEG dalgaları [40]

2.2.1.2 Elektrookülografi (EOG)

EOG iki temel amaç için kayıt edilir. 1.si uykunun başlangıcının teşhis edilmesini kolaylaştıran yavaş salınımlı göz hareketlerinin görüldüğü evre 1’i diğer evrelerden ayırt etmek içindir. 2.si ise hızlı göz hareketleri sayesinde REM evresini belirlemek için önem arz etmektedir [41]. Sağ göze ‘ROC’, sol göze de ‘LOC’ ismi verilir.

EOG derivasyonları için daha önceleri 2 sistem önerilmekteydi. Birincisi Rechtschaffen ve Kales’tir önerdiği 2 elektrot ta tek kulak elektrotu referans alınır (E1-A1, E2-A1 gibi). Mayo kliniğın önerdiği bağlama yöntemindeyse elektrotlar dış kantus lateralinin altına yerleştirilir ve referans noktası olarak Fpz alınır [30].

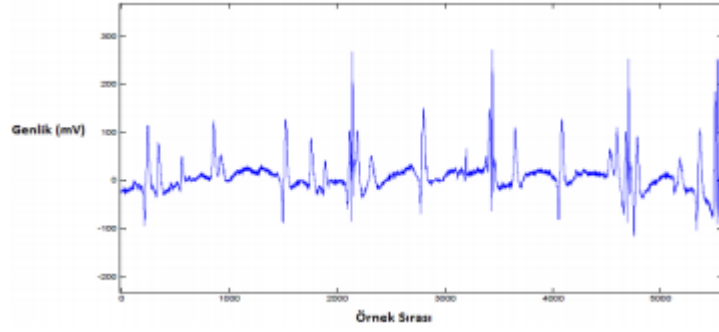
EOG gözün iç kısmında bulunan retina pigment epiteli ile fotoreseptör hücrelerinin dış segmentleri arasındaki dinlenim potansiyelidir. Normal gözde bu potansiyel 5 mV’tur ve ışık şiddetiyle değişir. Frekansları 0–100 Hz aralığında değişir [42].

2.2.1.3 Elektromyografi (EMG)

Standard polisomnografik incelemede EMG kaydı REM uykusunun tanınması amacıyla kullanılır. En yaygın olarak kullanılanı çene EMG'dir. EMG elektrotları yerleştirilirken elektrotların çene kemiğine değil kasa denk gelecek şekilde yerleştirilmesine dikkat edilmelidir [29].

EMG kaydında normal elektrot kullanılacaksa elektrotlar kol ekstensör kaslarına aralarında 2-3 cm olacak şekilde yerleştirilir. Eğer Piezo-elektrik bant kullanılıyorsa tek bir bant kullanıldığından kasın orta bölümüne yerleştirmek yeterli olur [35].

EMG işaretleri vücut üzerinde iğne veya yüzeysel elektrotlarla alınır. Bu tür işaretlerin kuvvetlendirilmesinde diferansiyel (fark) kuvvetlendiricisi kullanılır. İşaretin değerlendirilmesinde zaman domeninde ve/veya frekans domeninde yapılabilir. Aşağıda örnek bir EMG sinyali gösterilmiştir [39].



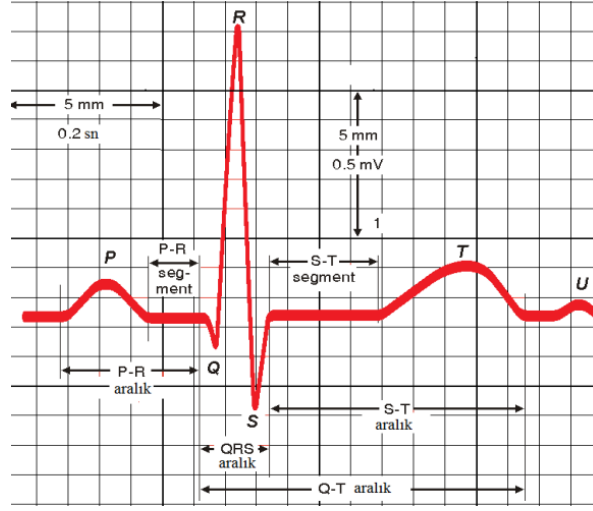
Şekil 2.3 Örnek bir EMG sinyali

2.2.1.4 Elektrokardiyografi (EKG)

EKG kaydında tek derivasyon kullanılır. Elektrotlar klasik derivasyona paralel şekilde sağ ve sol omuza yerleştirilir. Bu derivasyon kalp hızı, ritmi ve solunum olaylarının kaydedilmesini sağlar [29]. Standard EKG elektrotu kullanımı oldukça iyi sonuç verir [41].

Kalp kaslarının periyodik olarak kasılıp gevşemesi sonucu ortaya çıkan elektriksel işaret eğrilerine elektrokardiyogram denir [43]. Bu işaretlerin genlikleri en çok 2 mV, frekansları 0.05 Hz ile 100 Hz arasındadır. Normal EKG işareti kalbin dinlenme durumundaki taban seviyesi üzerinde sıralanan P, Q, R, S, T harfleri ile

sembolize edilen dalgalardan oluşur. Aşağıda Şekil 2.4’de EKG işaretinin bölümleri gösterilmiştir [44].



Şekil 2.4 EKG işaretinin bölümleri [44]

2.2.1.5 Hava Akımı

Hava akımı ölçümü direkt olarak pnömotakometreyle yapılabilir ancak bu yöntemde hastanın yüzüne tam oturan bir maske gereklidir. Bu yöntem pratik olmadığı için pek tercih edilmez. Bu yöntem yerine indirekt 2 yöntemden biri kullanılabilir. Birincisi; termokupl ve termistörlerin ağız ve her iki burun deliği önüne yerleştirilerek hava akımına sekonder ısı değişiminin ölçülmesidir. Termistör oda sıcaklığından etkilenebildiği için son dönemlerde tercih edilmemektedir [45]. Aşağıdaki Şekil 2.5’de termokupl ve termistörler ile hava akımı ölçme cihazları gösterilmiştir.



Şekil 2.5 Termokupl ve termistörler ile hava akımı ölçme cihazları [46]

İkinci yöntemse ‘Nazal Basınç Transdüser’dir. Hava akımına bağlı basınç değişikliklerinin burun deliklerinin önüne yerleştirilen nazal kanül ve basınç transdüser yardımıyla yorumlanması esasına dayanır. Hava akımı kısıtlamalarının saptanmasında çok etkilidir [47]. Uygun örneklem hızıyla kaydedildiğinde dalga formu değişiklikleriyle de solunum problemlerinin varlığı ve niteliği hakkında bilgi sağlar [29]. Aşağıda Şekil 2.6’da Nazal basınç transdüseri ile hava akımı ölçülmesi gösterilmiştir.



Şekil 2.6 Nazal basınç transdüseri ile hava akımı ölçülmesi [46]

2.2.1.6 Solunum Eforu

Solunum eforunun direkt ölçümü atmosfer basıncına göre gelişen basınç değişikliklerinin saptanmasıdır. Üç farklı yöntemle bu değişim saptanabilir;

Özofagus basınç ölçümü [41]

İndüktans pletismografi [48]

Piezoelektrik bantlar [41]

2.2.1.7 Oksihemoglobin Saturasyonu

Kan gazlarındaki değişiklikleri saptamak azalmış hava akımının oluşturduğu fizyolojik etkileri saptamayı sağlar. Oksihemoglobin saturasyonunu ölçmek için kullanılan en sık yöntem ‘pulse oksimetre’dir. Oksimetre, kan tırnak yatağı, kulak memesi gibi bölgelerden geçerken içerisinde bulunan farklı hemoglobin formlarının kırmızı ışığı absorbe etme düzeylerine bağlı olarak ölçüm yapar. Devamlı, dijital bir

veri akışı sağlar [29]. Aşağıda Şekil 2.7’de oksihemoglobin saturasyonu ölçme metodları gösterilmiştir.



Şekil 2.7 Oksihemoglobin saturasyonu ölçme metodları [41]

Bu cihazlar belli bir süre içinde yaptıkları ölçümü ortalayarak yansıtırlar. Bu süre 3 ila 21 saniye arasında değişir [41]. Dolaşım süresine bağlı olarak solunumsal olaydan 7 ila 9 saniye sonra kulak memesinde veya biraz daha uzun süre sonra tırnak yatağında oksihemoglobin saturasyon grafiği en düşük düzeyine ulaşır [30].

2.2.1.8 Pozisyon

Pozisyon sensörü civalı bir mekanizmadır. Bu sensör hastanın yatış pozisyonu hakkında bilgi vermektedir. Göğüs ortasına yerleştirilen sensörle hastanın yatış pozisyonu ve bu pozisyonda görülen apne-hipopne sıklığını hesaplamak için kullanılmaktadır [41]. Aşağıdaki Şekil 2.8’de pozisyon sensörü gösterilmiştir.



Şekil 2.8 Pozisyon sensörü [41]

2.2.1.9 Horlama

Horlama gibi sesler mikrofon yardımıyla kaydedilir. Çıktı direkt olarak poligraf üzerinde gösterilebileceği gibi entegre sinyal olarak da alınabilir [41]. Horlama mikrofon yardımıyla kayıt edilmektedir.



Şekil 2.9 Horlama ölçümünde mikrofon kullanılması [41]

2.2.1.10 Karbondioksit Ölçümü

Uyku apne sendromlu hastalarda gece boyunca oluşan apne ve hipopneler sırasında saturasyonda düşmeler olabileceği gibi karbondioksit seviyesinde artış da olabilir. Kızıl ötesi ısınları kullanan sensör yardımıyla CO₂ ölçülür [29].

2.2.1.11 Özofageal Ph Monitörizasyonu

Gece gastroözofageal reflüden yakınan hastalara uygulanır. Hastalarda gece boyunca oluşan reflü takip edilebildiği gibi bu reflü ataklarının apne, hipopne ve arousallarla ilişkisi olup olmadığı tespit edilebilmektedir [35].

2.2.1.12 Noktürnal Penil Tümesans ve Vücut Isısı Takibi

Eretil disfonksiyon şüphesi olanlarda gece boyunca penis çapındaki değişimin ölçülmesidir. Vücut ısı takibi sirkadyen ritim bozuklukları tanısında kullanılır [34].

2.3 UPHB Tanı Kriterleri

UPHB polysomnografi kayıtlarına bakılarak teşhis edilmektedir. Hareketler 0,5 sn. ile 5 sn. arası sürmelidir. Minimum 8 mikrovolt kadar dinlenme EMG'sinin genliğinin üstünde olmalı. Ardı ardına 4 ya da daha fazla hareket olmalıdır [49]. Hareketler arası interval (bir hareketin bağlamasından diğerinin başlamasına kadar geçen süre 5 sn. den uzun ve 90 sn.den kısa) olmalıdır (tipik olarak 20–40 sn. olmalıdır). UPHB indeksi çocuklarda 5/saat, yetişkinlerde 15/saat in üzerindedir [50].

2.4 Veritabanı

Veritabanı birbiriyle ilişkili verilerin tekrar etmeden, çok amaçlı kullanımına uygun şekilde depolanmasıdır. Birbirleriyle ilişkileri olan veya olmayan verilerin tutulduğu, numerik verilerin ve fiziksel olarak tanımlarının olduğu depolama alanlarıdır [51].

Belirli bir konu hakkında toplanmış veriler bir veritabanı programı altında toplanır. İstenildiğinde toplanan bilgilerin tümü veya istenilen özelliklere uyanları görüntülenebilir, yazdırılabilir hatta bilgilerden yeni bilgiler üretilerek bunlar çeşitli amaçlarla kullanılabilir [51].

Farklı amaçlar için veri miktarı, işlem sayısı, güvenlik ve buna benzer farklı özelliklerde birçok veritabanı yönetim sistemleri bulunmaktadır. Yapılan çalışmada amacımıza uygun ve ücretsiz olan MySQL seçilmiştir.

MySQL, hızlı, sağlam güvenilir bir veritabanı yönetim sistemidir. MySQL veritabanı ile tabloların kontrolü, tabloların düzenlenmesi, optimizasyonu ve tamiri çok hızlı olarak yapılabilmektedir [51].

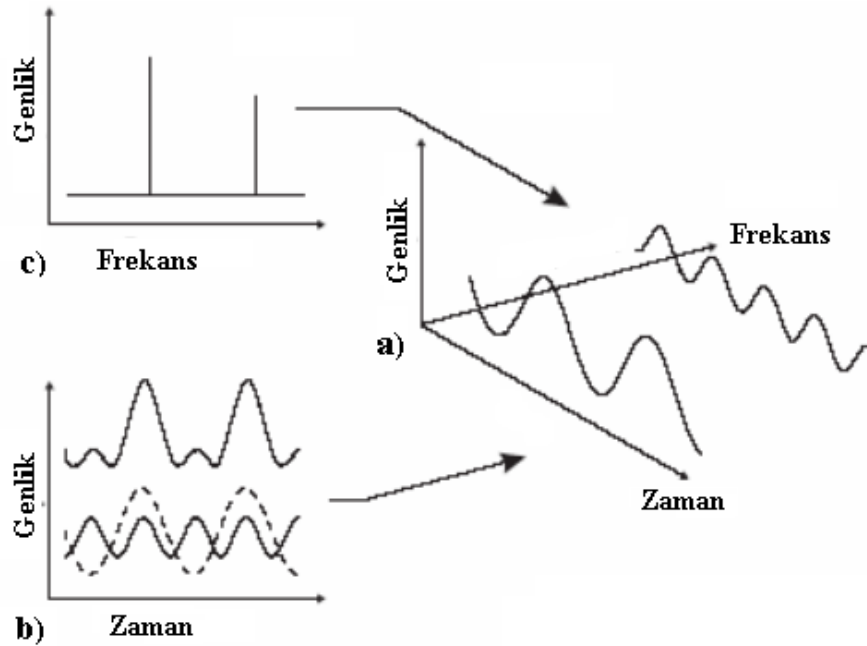
UNIX ve OS/2 platformları için ücretsiz dağıtılmaktadır. MySQL açık kaynak kodulu bir veritabanıdır. Geliştiricileri tarafından, 500'den fazlası 7 Milyon kayıt içeren 10,000 tablodan oluşan kendi veritabanlarını neredeyse 100 gigabyte civarında veriyi MySQL'de tutabilmektedir [51].

2.5 Çalışmada Kullanılan Dijital Sinyal İşleme Metotları

2.5.1 Fourier Analiz

Tarihsel olarak sinyallerin analiz yöntemleri temeli 19.yy'a kadar uzanmaktadır [52]. Bu yüzyılda J. Fourier herhangi bir periyodik fonksiyonun, sonsuz sayıda periyodik-karmaşık-üstel fonksiyonun (sinüzoidlerin) toplamı şeklinde tanımlanabileceğini göstermiştir. J.Fourier'in fikirleri yıllar sonra önce periyodik olmayan fonksiyonlar için daha sonra da hem periyodik hem de periyodik olmayan ayrık zaman sinyalleri için uygulanabilir hale gelmiştir. Böylece bu işlemler bilgisayarlar için de uygun hale gelmiştir [53].

Fourier'in de ispatladığı gibi ve şu anda da çok kullanılan bir veri sinyali, sinüs sinyallerinin genlik, frekans ve faz değerleri doğru şekilde seçilerek birleştirilmesi ile elde edilebilir. Bir veri sinyali birden fazla sinüs sinyaline ayrıştırılabilir [6].



Şekil 2.10 Zaman-Frekans-Genlik düzlemi [54]

Sinyaller, sinüs sinyallerinin zaman-genlik veya zaman-frekans bölgelerinde Şekil 2.10'daki gibi gösterilebilir. Sinyallerin frekans bölgesindeki gösterimine sinyal spektrumu denir. Her bir spektrumdaki çizgi ise sinyalin bileşeni olarak adlandırılır.

Frekans spektrumunu sinyali ifade etmek için kullanırız. Frekans spektrumuna, sinyalin frekans bilgisinin, grafiksel gösterimi de diyebiliriz [6].

2.5.1.1 Kısa-Zaman Fourier Dönüşümü

Fourier Dönüşümü, f frekanslı bileşen zamanın hangi anında ortaya çıkarsa çıksın integrasyona etkisi aynı olacaktır. f frekanslı bileşenin t_1 ya da t_2 anında ortaya çıkması integrasyon sonucunu değiştirmeyecektir. Fourier Dönüşümü, yalnızca belirli bir frekans bileşeninin var olup olmadığını belirtmektedir (Fourier Dönüşümü ile işaretin sadece spektral içeriği elde edilir) [11].

Kısa Süreli Fourier Dönüşümü'nde durağan olmayan işaret, zamanda durağan kabul edilebilecek küçük parçalara bölünür. Diğer bir deyişle Fourier Dönüşümünden farklı olarak, işarete dar pencerelerden bakılır ve pencere içinde kalan işaretin durağan olduğu varsayılır. Aşağıdaki denklem ile gösterilmiştir. Kısa Süreli Fourier Dönüşümünün ifadesi görülmektedir [11].

$$KZFD(\tau, f) = \int_{-\infty}^{\infty} [x(t) \cdot w^*(t - \tau)] \cdot e^{-j2\pi ft} \cdot dt \quad (2-1)$$

$x(t)$, orijinal işareti; $w(t)$, pencere fonksiyonunu ve $*$, karmaşık eşleniği göstermektedir. f , frekans; τ ise zamanda öteleme miktarıdır [11].

Denklemden görüldüğü gibi Kısa Süreli Fourier Dönüşümü, bir pencere fonksiyonu ile çarpılan $x(t)$ 'nin Fourier Dönüşümü'nden başka bir şey değildir [11].

Her τ ve f için yeni bir Kısa Süreli Fourier Dönüşümü katsayısı hesaplanır. Fourier Dönüşümü sadece frekansın bir fonksiyonu iken, Kısa Süreli Fourier Dönüşümü hem frekansın hem de zamanın bir fonksiyonudur ve dönüşüm bu haliyle iki boyutludur [11].

İşaretin zaman-frekans temsili elde edilmesine rağmen, seçilen pencerenin genişliği dönüşümün etkinliğinde önemli rol oynamaktadır. Kısa Süreli Fourier Dönüşümünde pencere genişliği ile ilişkili bir çözünürlük problemi bulunmaktadır [11].

Fourier Dönüşümü'nde frekans domeninde çözünürlük problemiyle karşılaşılmaz. Çünkü hangi frekansların var olduğu kesin olarak bilinir. Kısa Süreli Fourier Dönüşümü'nde seçilen pencerenin sonlu uzunlukta olması nedeniyle işaretin bir parçası ele alınır ve frekansta çözünürlük kötüleşir. Frekans çözünürlüğünün artırılması

uğruna pencerenin geniş tutulması ise belirli bir frekans bandının hangi zaman aralığında ortaya çıktığını belirsiz kılar, yani zamanda çözünürlüğü kötüleştirir [55].

Pencere genişliği durağanlık varsayımını geçerli kılacak kadar dar olmalıdır. Dar bir pencere seçilmesi durumunda hem bu varsayım geçerliliğini koruyacak hem de Fourier Dönüşümünde sağlanamayan zamanda çözünürlük iyileşecektir. Pencere genişledikçe frekans çözünürlüğü artar; ancak zamanda çözünürlük azalır. Sonuçta işarete Kısa Süreli Fourier Dönüşümünü uygulamadan önce zamanda ya da frekansta çözünürlüğün sağlanması ikilemiyle karşılaşılmaktadır [56].

Kısa Süreli Fourier Dönüşümü, tüm zamanlarda sabit çözünürlük verdiği için Kısa Süreli Fourier Dönüşümünün çözünürlük ile ilgili problemlerini gidermek üzere zamanda değişken çözünürlük Dalgacık Dönüşümü geliştirilmiştir [57]. Dalgacık Dönüşümüne frekans cevabı zamanla değişen durağan olmayan işaretlerin analizinde ihtiyaç duyulmaktadır.

2.5.1.2 Ayırık Zamanlı Fourier

Ayrık zamanlı Fourier dönüşümü bir ayırık zaman işaretini sürekli frekans bileşenlerine ayırmaktadır. Bu nedenle bir ayırık zamanlı işaretin ayırık zamanlı Fourier dönüşümü frekansa bağlı sürekli bir fonksiyondur [58].

Sayısal sinyal işlemenin pratik uygulamalarının çoğunda, bilgisayar sonsuz $x(n)$ dizisinin elemanlarını saklayamaz ve sürekli w frekansının değerlendirilmesi olanaksızdır. Ayrıca teorik olarak tanımlanan bazı serilerin aksine gerçek serilerin Fourier Dönüşüm'leri hesaplanamamaktadır. Bu nedenle sayısal sinyaller için Fourier dönüşümünün kullanılması uygun olmamaktadır. N uzunluğunda, (sonlu) bir $x(t)$ ayırık sinyal için 'Ayırık Fourier Dönüşümü' (AFD) [58].

$$X_s(f) = \sum_{t=0}^{N-1} x(t) \cdot e^{-j f \frac{2\pi t}{N}} \quad (2-2)$$

$$x(t) = \frac{1}{N} \sum_{f=0}^{N-1} X_s(f) \cdot e^{j f \frac{2\pi t}{N}} \quad (2-3)$$

eşitlikleriyle tanımlanmaktadır. AFD'nin doğrudan hesaplanmasında her bir $X_s(f)$ değeri için N karmaşık çarpma ve $N-1$ karmaşık toplama işlemi kullanılmaktadır. Bu durumda N adet AFD (Ayırık zamanlı Fourier dönüşümü) değeri hesaplanırken, N^2 çarpma ve $N(N-1)$ toplama işlemi yapılmaktadır [58].

AFD doğasındaki periyodiklik nedeniyle bir sinyalin N-noktalı AFD alındığında işaretin N periyodu ile periyodikmiş gibi işlem görmektedir. Bu nedenle, AFD hesabı, en az sinyalin örnek sayısı kadar ayırık frekans değerinde gerçekleştirilmesi gerekmektedir. AFD sinyalin örnek sayısından daha az sayıda ayırık frekans değerinde hesaplandığında, sinyalin frekans spektrumunun seyrek örneklenmesi nedeniyle zamanda örtüşme (aliasing) meydana gelmektedir. Bu durumda zaman örtüşmesinden dolayı işaret değerleri Ters-AFD ile geri oluşturulamamaktadır. AFD hesabındaki N değeri sinyalin örnek sayısından fazla olabilir ve bu durumda işaretin sonuna sıfır değerlerinin eklenmesi ile sinyalin uzunluğu N'ye çıkarılabilmektedir [58].

AFD hesaplamasında etkin ve günümüzde kullanılan yaklaşım HFD algoritmalarıdır. 1965 yılında Cooley ve Tukey tarafından kurulan HFD algoritması; AFD'den farklı değildir. AFD'nin hesaplanması için etkili ve mükemmel bir algoritmadır. AFD'nin sayısal sinyal işleme alanında spektrum analizi ve korelasyon gibi işlemlerin yapılmasında önemli rol oynamasının nedeni HFD algoritmalarından kaynaklanmaktadır [59].

Bir sinyalin frekans spektrumunun hesapsal yöntemlerle elde edilmesi için 2π ile periyodik olan ayırık-zamanlı frekansın bir periyodunun dikkate alınması yeterli olmaktadır. Ayırık zamanlı frekans spektrumunun 2π 'lik temel periyodunda N adet eşit aralıklı frekans değeri;

$$\omega = \frac{2\pi}{N}k, \quad k = 0,1,2, \dots, N - 1 \quad (2-4)$$

şeklindedir. Burada k tamsayısı 2π 'lik temel frekans bandındaki ayırık frekans değerlerini belirtmektedir ve frekans endeksi olarak adlandırılmaktadır. Fourier dönüşümü için N adet ayırık frekans değeri için hesaplandığından, $k=0,1,\dots,N-1$, toplam N adet karmaşık Fourier Dönüşüm değeri $X_s(f)$ elde edilmektedir. Bir sinyalin frekans spektrumu için sinyalin genlik, faz ve güç spektrumları HFD sonucu elde edilen karmaşık sayılardan kolayca belirlenebilmektedir [60].

2.5.1.3 Akan Fourier Dönüşümü

Bir $f(t)$ sinyalinin Akan Fourier Dönüşümü ifadesi; c verilen bir sabit olmak üzere aşağıdaki ilk eşitlik ile verilir. $F(t,\tau)$, $f(t)$ sinyalinin dilimine karşılık gelen $f(t+\tau)$

sinyalinin Fourier dönüşümüdür. Böylelikle durağan olmayan sinyaller için spektrumun değişimi sinyalin durağan kabul edildiği dilimleri ile belirlenebilmektedir [61].

$$F(t, \omega) = \int_{-c}^c f(t + \tau) e^{-j\omega\tau} d\tau = e^{j\omega} \int_{t-c}^{t+c} f(\alpha) e^{-j\omega} d\alpha \quad (2-5)$$

$F(t, \omega)$ nin Ters Dönüşümü ifadesi de $\omega_0 = \frac{\pi}{c}$ olmak üzere;

$$f(t) = \frac{1}{2c} \sum_{m=-\infty}^{\infty} F(t, m\omega_0) \quad (2-6)$$

Eşitliği ile verilir.

2.5.1.4 Hızlı Fourier Dönüşümü

Ayrık Fourier Dönüşümünün doğrudan hesaplanmasında her bir $f[k]$ değeri için N karmaşık çarpma ve $N-1$ karmaşık toplama işlemi kullanılmaktadır. Bu nedenle N adet AFD değeri bulunurken, N^2 çarpma ve $N(N-1)$ toplama işlemi gereklidir. Ayrıca her karmaşık çarpma işlemi için dört gerçel çarpma ve iki gerçel toplama işlemi ve her bir karmaşık toplama iki gerçel toplama işlemi ile gerçekleştirilmektedir. Neticede, dizi uzunluğu olan N 'nin 1000'in üzerinde olması halinde doğrudan AFD' nin hesaplanması çok fazla miktarda işlem gerektirmektedir. Yani N sayısı artarken gereken işlem sayısı çok fazla artmaktadır [62].

Son olarak Ayrık zamanlı bir $f[k]$ işaretinin $w[k]$ ayrık pencere dizisi kullanılarak AFD ifadesini aşağıdaki gibi

$$f(t) = \frac{1}{2c} \sum_{m=-\infty}^{\infty} F(t, m\omega_0) \quad (2-7)$$

ile verebiliriz.

$$F[n] = AFD\{f[k]\} = \sum_{k=0}^{N-1} f[k]w[k]e^{-j2\pi kn/N} \quad (2-8)$$

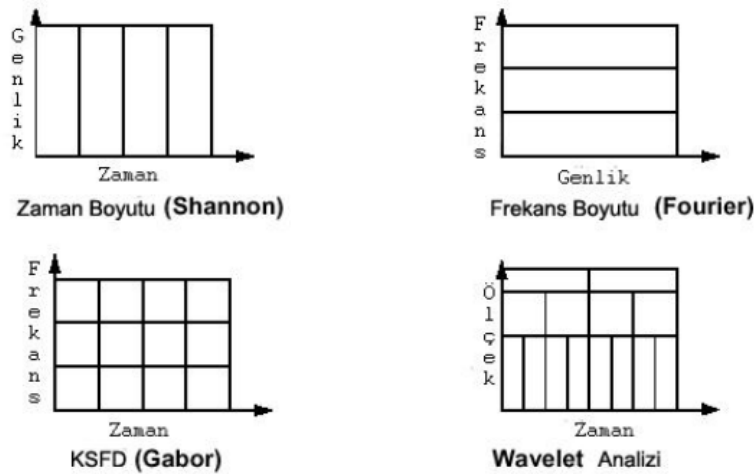
2.5.2 Dalgacık Dönüşümü

Dalgacık dönüşümü Fourier dönüşümünün durağan olmayan sinyallerdeki eksiklerini gidermek amacıyla ortaya atılmıştır. Bu analiz yöntemi gürültüye karşı daha az hassasiyet göstermekte ve durağan olmayan sinyallere rahatlıkla uygulanabilmektedir [63]. Günümüzde hem biyolojik işaretlerin (EEG, EMG, EKG) analizinde hem de hidrolojik çalışmalarda su akımlarının incelenmesi [64], günümüz

otomasyon sistemlerinde kullanılan rulmanların hata tespit algoritmalarının karşılaştırılmaları [65], tekstil imgelerinde hata sezimi [66] ve elektrik motorlarında arıza tanılarının konulması [67] gibi çok farklı alanlarda kullanılmaktadır.

Dalgacık dönüşümünde analiz edilen sinyal, dalgacık olarak adlandırılan yapılar cinsinden temsil edilir. Dalgacıklar, çevirme ve ölçekleme işlemleriyle ana dalgacık olarak adlandırılan tek bir fonksiyondan türetilen bir fonksiyonlar ailesidir. Ölçekleme işlemi, ana dalgacığın daralması veya genişlemesi olarak bilinir [68]. Çevirme işlemi ise zaman ekseninde ana dalgacığın yerinin değişmesidir [69].

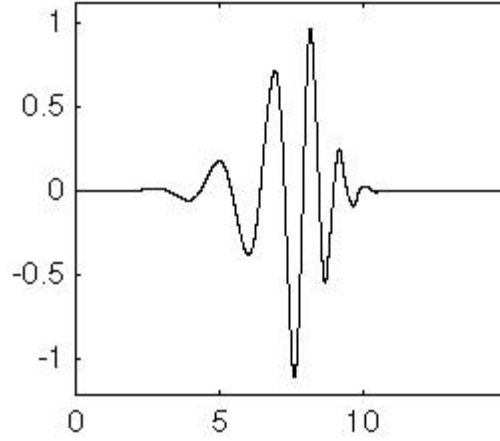
Dalgacık Dönüşümü, değişken büyüklüğünde pencereleme ile işaret işleme yöntemlerine yeni bir yaklaşım getirir. Düşük frekanslarda uzun zaman aralıklı, yüksek frekanslarda ise kısa zaman aralıklı pencereleme yapılmasını sağlar. Aşağıda KSFD ile Dalgacık Dönüşümü arasındaki farklılık şekil üzerinde gösterilmektedir (Şekil 2.11).



Şekil 2.11 KSFD ile Dalgacık Dönüşümü arasındaki farklılıklar [60]

Dalgacık Dönüşümü, zaman-frekans bölgesi yerine zaman-ölçek bölgesini kullanır. Dalgacık Dönüşümü'nün en önemli avantajı, işaretin lokal olarak analizine imkan sağlamasıdır. Örneğin, kısa süreksizliğe sahip bir sinüzoidal bir dalga ele alınırsa, dalgacık katsayıları süreksizliğin zaman boyutunda tam olarak nerede olduğunu verebilir. FD ise bunu gösteremez. Ayrıca Dalgacık Dönüşümü, sıkıştırma ve gürültüden temizleme işlemlerinde de başarılı sonuçlar verir [70].

Sinüs dalgası, sürekli değildir, eksi sonsuz ile artı sonsuz arasında sonsuz salınır, düzgün ve tahmin edilebilirdir. Dalgacık ise düzgün değildir ve asimetriktir [70].

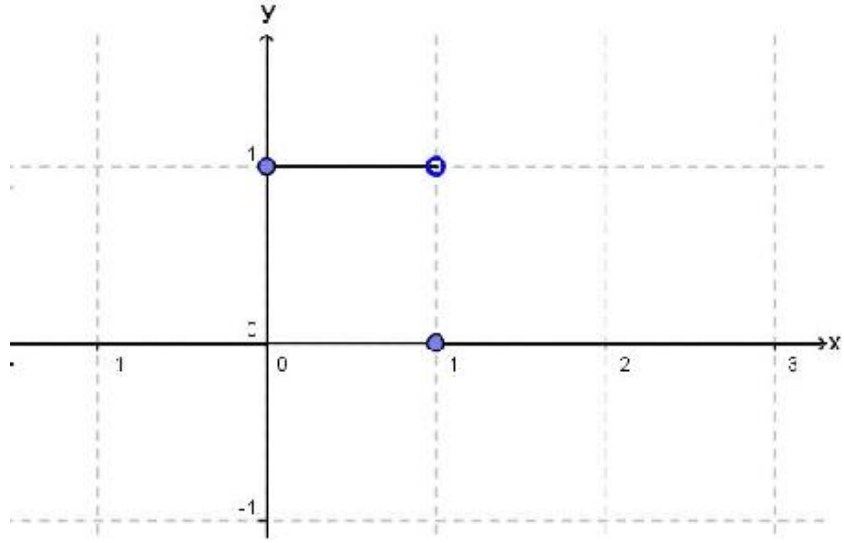


Şekil 2.12 Dalgacık (Daubechies) [60]

Fourier analizi, bir işaretin farklı frekansları için dağıtılmış sinüsler toplamından oluşur. Benzer bir şekilde, dalgacık analizi de orijinal dalgacığın kaydırılmış, ölçeklenmiş halinin işaret içine dağıtılmış halidir. Yalnızca sinüs ve dalgacık işaretlerine bakarak bile, ani değişimlere sahip bir işareti analiz etmede, asimetrik dalgacığın daha iyi sonuçlar vereceği düşünülebilir. Ayrıca lokal durumlar için de dalgacık işaretinin daha başarılı olabileceği görülebilir [70].

2.5.2.1 Haar Dalgacık Dönüşümü

Çeşitli dalgacık dönüşüm tabanları mevcut olmakla birlikte, Haar dalgacık dönüşümü sinyal karakteristiklerinin zaman alanındaki konumlarını ortaya koyma konusunda en basit ve kısa dönüşüm yoludur [71]. Bu yüzden bu çalışmada kullanılan dalgacık dönüşümlerinde Şekil 2.13’de grafiksel olarak gösterilen Haar dalgacığı ana dalgacık olarak seçilmiştir.



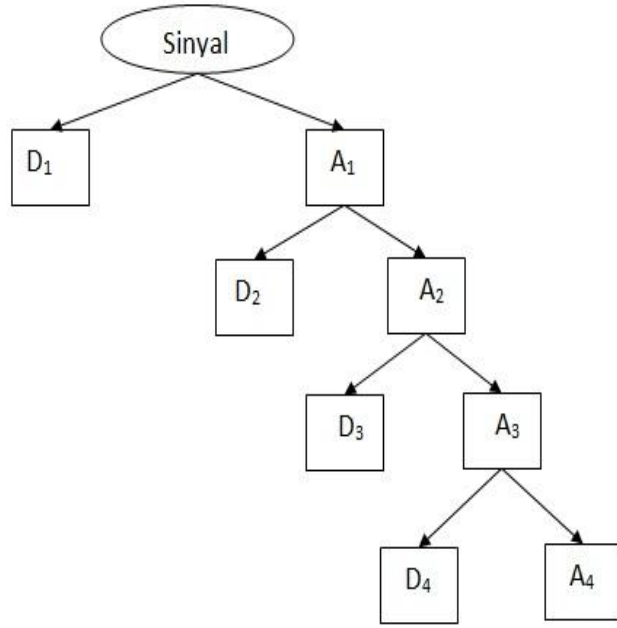
Şekil 2.13 Haar dalgacığının grafiksel gösterimi

Sinyali bileşenlerine ayırıştırma işlemi için yüksek geçiren bir filtre, $(g(n))$ ve alçak geçiren bir filtre $(h(n))$ kullanılır. Bu filtreler Haar dalgacığından türetilmiştir ve matematiksel ifadeleri aşağıdaki eşitliklerde gösterilmiştir.

$$g(n) = \left\{ \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right\} \quad (2-9)$$

$$h(n) = \left\{ \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\} \quad (2-10)$$

Alçak geçiren ve yüksek geçiren filtreler, 2^1 . seviye dalgacık dönüşümünü gerçekleştirmek için giriş sinyali olarak analizi yapılan sinyali alır ve çıkışında yaklaşım katsayıları (A_1) ve detaylı (D_1) katsayılardan oluşan, örneklerinin yarısı atılmış bir sinyal üretir. 2^2 . derece dalgacık dönüşümü için ise 2^1 . derece dalgacık dönüşümü sonucu elde edilmiş yaklaşım (A_1) katsayıları kullanılarak 2. Seviyeye ait yaklaşım katsayıları (A_2) ve detaylı (D_2) katsayılar üretilir. Dalgacık dönüşümünün seviyesi artırıldıkça her seferinde bir önceki seviyenin yaklaşım (A_n) katsayıları kullanılarak aynı işlem tekrar edilir (Bkz.Şekil 2.14) [72].



Şekil 2.14 Dalgacık Dönüşümü'nün seviyesi artırıldıkça, yaklaşım katsayıları ve detaylı katsayıların nasıl üretildiğini gösteren blok diyagram

2.6 Makine Öğrenmesi Nedir?

Makine öğrenmesi, bilgisayarların geçmiş bilgilerden elde edilen tecrübelerden yararlanarak, gelecekteki olayları tahmin etmesine ve modelleme yapmasına imkan veren bir yapay zeka alanıdır. Bilgisayarın bir olay ile ilgili bilgileri ve tecrübeleri öğrenerek gelecekte oluşacak benzeri olaylar hakkında kararlar verebilmesi ve problemlere çözümler üretebilmesi olarak da tanımlanabilir. Makine öğrenimi araştırmalarının odaklandığı konu bilgisayarlara karmaşık örüntüleri algılama ve veriye dayalı akılcı kararlar verebilme becerisi kazandırmaktır. Bu, makine öğreniminin istatistik, olasılık kuramı, veri madenciliği, örüntü tanıma, yapay zeka, uyarlamalı denetim ve kuramsal bilgisayar bilimi gibi alanlarla yakından ilintili olduğunu göstermektedir [73].

Makine öğrenmesi ile bilgisayar yazılımlarının, olayları daha önceki örneklerden ediniliş tecrübelerin kullanılarak öğrenmesi sağlanır. Öğrenecek olan bilgisayar

sistemleri, önce bir örnek almakta ve bu örnekten bazı bilgileri öğrenmektedir. Daha sonra, ikinci örneğe bakarak biraz daha bilgi edinmektedir. Bu işlemi öğrenilecek olay ile ilgili genellemeler yapılmaktadır. Bu olaya tecrübelerden öğrenmenin bir yolu olarak bakmak mümkündür [74].

Makine öğrenmesi ve veri madenciliği arasında doğrudan bir ilişki vardır. Makine öğrenmesi tekniklerinin büyük veritabanlarına uygulanması veri madenciliğidir [75]. Makine öğrenmesi veri madenciliği sürecinde uygulama aşamasında yer almaktadır. Bu aşamada seçilen bir makine öğrenmesi tekniği veri seti üzerine uygulanarak sonuçlar elde edilir. Makine öğrenmesi yalnızca veriler üzerinde uygulanan bir teknik değil, bir yapay zeka alanıdır. Veri madenciliği çıkan bilgi ve bu bilgilerin değerlendirilmesi ile uğraşmaktadır. Makine öğrenmesi ise bu bilgiyi çıkarmayı sağlayan teknikler ve bu teknikleri kullanan bilgisayar programlarının kendilerini geliştirmesi ile ilgilenir. İki konu arasındaki en büyük fark; makine öğrenmesi, öğrenme metotlarını geliştirerek, tahminleri ya da tanımları en iyi şekilde, yüksek performans ile nasıl çıkarabileceği ile ilgilenirken, veri madenciliğinin ortaya çıkan bilgi ile ilgilenmesidir [76].

Makine öğrenmesinin günlük hayatımızdaki bazı kullanım alanları aşağıda verilmiştir [73];

- El yazısı veya imza tanıma sistemleri [73]
- Tıbbi verileri tanımlamada kullanılan sistemler [73]
- Metin ve Mail analizinde kullanılan sistemler [73]
- Parmak izi tanıma sistemleri [73]
- Göz taraması (İris) ile tanıma sistemleri [73]
- Yüz tanıma sistemleri [73]

Bilgi teknolojilerindeki gelişmeler sayesinde, artık çok büyük miktardaki veriyi kaydedebilmekteyiz. Makine öğrenmesi metotları geçmişteki veriler kullanılarak veriye en uygun modeli bulmaya çalışırlar. Yeni gelen verileri de bu modele göre analiz ederler. Büyük miktarda verinin incelenip onun içinden işe yarayan bilginin (modelin) elde edilmesi işlemine, veri madenciliği (data mining) de denilmektedir. Farklı uygulamaların analizlerinden farklı beklentileri olmaktadır. Makine öğrenmesi metotlarını bu beklentilere göre sınıflandırmak mümkündür [75].

1.Sınıflandırma: Geçmiş bilgilerin hangi sınıflara ait olduğu verildiğinde yeni gelen verinin hangi sınıfa dahil olduğunun bulunması işlemidir [19].

2.Kümeleme: Geçmiş bilgilerin sınıflarının/etiketlerinin verilmediği/bilinmediği durumlarda verilerden birbirine benzerlerin yer aldığı kümelerin bulunması işlemidir [19].

3.Eğri uydurma (Regresyon): Geçmiş verilerin sınıflarının sürekli sayılar olduğu durumlarda kullanılır. Örneğin bir hisse senedinin değeri bir sayıdır ve bu senede ait model bu değeri tahmin etmeye yönelik bir eğri uydurma işlemi olacaktır [19].

4.Özellik seçimi/Çıkarımı: Veriye ait birçok özellikten, verinin kümesini/sınıfını/değerini belirleyen özelliklerinin hangileri olduğu bilinmeyebilir. Bu durumda tüm özellik kümesinin bir alt kümesi seçilir (özellik seçimi) ya da bu özelliklerin birleşimlerinden yeni özellikler elde edilir (özellik çıkarımı) [19].

5.İlişki Belirleme: Bir süper markette X ürününü alan müşterilerden %80'i Y ürününü de alıyorsa, X ürününü alıp Y ürününü almayan müşteriler, Y ürününün potansiyel müşterileridir. Müşterilerin sepet bilgilerinin (bir alışverişte alınan ürün bilgileri) bulunduğu bir veritabanında potansiyel Y müşterilerini bulma işlemi türündeki bulunduğu bir veritabanında potansiyel Y müşterilerini bulma işlemi türündeki problemler ilişki belirleme metotlarıyla çözülmektedir [19].

Makineler insanlığın işgücüne sağladıkları katkıyı, makine öğrenme metotları sayesinde insanlığın beyin gücüne de sağlamaya başlamışlardır. Her tür uygulama için çok miktarda verinin analiz edilerek gelecekle ilgili varsayımlar geliştirmemize, kararlar vermemize yardımcı olan makine öğrenme metotları önemleri ve katkıları artmaktadır [19].

2.6.1 Makine Öğrenmesinin Amaçları

Makine öğrenmesinde, tümevarım kullanılarak çıkarımlar yapılmaktadır. Yapılan çıkarımlar, geleceğe yönelik tahminde bulunmak ya da bir tanım yapmak için kullanılır. Veri, geleceğe yönelik bilgi tahmini için kullanılacaksa, yani [75]'de tanımlandığı gibi oluşturulacak model tahmin edici model ise, iki aşama gereklidir: eğitim aşaması ve test aşaması [75].

Eđitim ařamasında, belirli miktarda veri kullanılarak bir model oluřturulur. Kullanılan veri, eđitim verisi olarak adlandırılır. Eđitim verisinin ne kadar ve nasıl seileceđi ayrı bir konudur. Oluřturulan model, sadece rnek veriyi deđil tm veriyi temsil eder. Test ařamasında ise, eđitim ařaması sonucunda oluřan modele, test ařaması iin ayrılan ya da ileride toplanacak olan veriler sunulur. Ortaya ıkan bilgi ise tahmin etmek iin kullanılır [75].

Makine đrenmesinin amalarını Mitchell, Michalski ve Carbonell makalelerinde  aıdan incelemiřlerdir [77].

1. Hedef-tabanlı alıřmalar: đrenme sisteminin geliřimi ve analizi, belirlenmiř grevleri yerine getirmek iin gerekleřtirilir. Bu yaklařım, makine đrenmesine mhendis yaklařımı olarak tanımlanmıřtır [77].

2. Biliřsel simlasyon: İnsanın đrenme srecini arařtırıp, bilgisayarda simlasyonunu gerekleřtirmek olarak tanımlanmıřtır. Bu ise, makine đrenmesine biliřsel modelleme yaklařımıdır [77].

3. Teorik analiz: Uygulama alanlarından bađımsız olarak teorik olabilecek đrenme metodları ve algoritmaları incelemek iindir [77].

Makine đrenmesi konusu ile uđrařan bilim adamları bu  yaklařımdan birini ya da her birini birbiri ile bađlantılı řekilde kullanarak alıřmalarını srdrmektedir [76].

Sunulan  yaklařımdan ilki olan mhendis yaklařımından yola ıkarak, Nilsson makine đrenmesinin nemini sıralamıřtır [78]. Bunlardan bazıları řunlardır:

- Bazı grevler iin, girdi/ıktıyı belirlesek de arasındaki iliřkiyi belirtilemeyebilir. Bu gibi durumda, makinenin kendi iyapısını ayarlayarak, byk veri yıđınlarından giriř/ ıkıř fonksiyonunu bulup, arasındaki iliřkiyi tahmin etmesi beklenir [78].

- Byk veri yıđını iinde gizli kalmıř nemli iliřkiler ve bađlantılar olabilir. Makine đrenmesi metodları, bu iliřkileri seip ıkarmak iin kullanılır. Bu konu, veri madenciliđi olarak adlandırılmaktadır [78].

- Bazı grevler iin bilgi miktarı, insanın kodlaması iin fazla olabilir. Bu gibi durumda makine insanın yapabileceđinden fazlasını yapabilir [78].

- Makine, deđiřikliklere kolayca adapte olabilir [78].

- Bazı görevlerde bilgiler deęişebilir. Bu gibi durumlarda, yapay zeka istemini tekrar tasarlamak pratik deęildir. Makine öğrenmesi metodlarını kullanarak bu gibi deęişiklikler gözlenebilir [78].

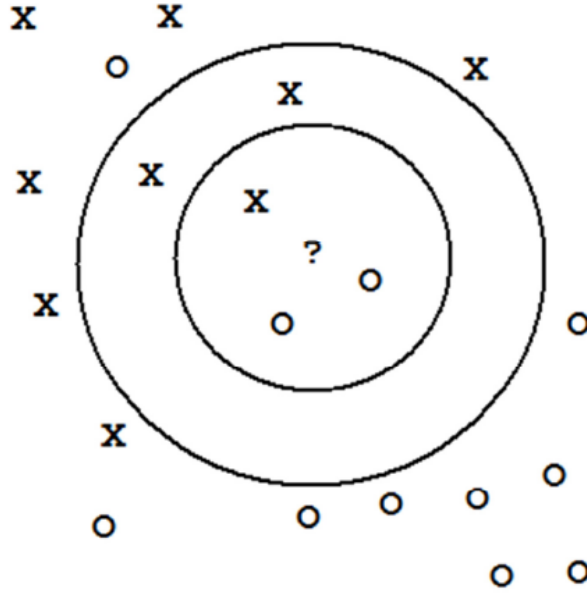
2.7 Makine Öğrenmesi Algoritmaları

2.7.1 En Yakın K-Komşu Algoritması

En yakın k-komşu algoritması sınıflandırma yöntemlerinden birisidir. Amaç, sınıfları belirli olan örnek kümeğe ait gözlemlere yeni bir gözlem eklendiğinde bu gözlemin hangi sınıfa ait olacağıın belirlenmesidir. Tanımdan da anlaşılacağı üzere geçmiş verilerin varlığına dayalı olup, gürültülü verinin varlığında sorun yaratır [79].

Bu algoritma sınıflandırma işlemini için gözlemler arası uzaklık hesabına dayanır. Son eklenen gözleme var olan gözlemlerin uzaklıklarının hesaplanması ve en az uzaklığı sahip k tane gözlemin seçilmesi işlemini gerçekleştirilir. Uzaklık hesabı için Öklid uzaklık formülü kullanılır [80].

KNN örnek tabanlı bir sınıflandırıcı olup sınıflandırıcı sistem parametreleri olarak sadece örnekler tutulmaktadır. Bu algoritma tüm örneklerin d boyutlu uzaydaki noktalar olduğunu kabul eder [81]. Sınıfı bilinmeyen yeni bir örnek test edildiğinde, test edilen örneğin kendi komsularından k tanesinin çoğunluğunun bulunduğu sınıfa atanır (Şekil 2.15). Bu şekilde '?' ile gösterilen örneğin 'X' mi yoksa 'O' grubundan mı olduğu bulunmak istenmektedir. Eğer k=3 seçilirse, 2 tane 'O' olduğu için test örneğinin 'O' sınıfından olduğu, fakat eğer k=5 seçilirse 3 tane X olduğu için test örneğinin 'X' sınıfından olduğu bulunacaktır [82]. k genellikle küçük değerli bir pozitif tam sayıdır. İki sınıflı problemlerde, k değerini tek sayı olarak seçmek sistemin test edilen örneğinin sınıfını belirlemede kararsız kalmasını engelleyecektir [83].



Şekil 2.15 KNN sınıflandırıcı örneği [84]

Bu sınıflandırıcı algoritmasında doğrudan önceden sınıfı bilinen örneklerle kıyaslama yapıldığından, ayrıca bir eğitim aşaması yoktur. Diğer bir deyişle, sınıflandırıcının eğitim aşaması sadece örneklerin ve ait oldukları sınıf bilgilerinin saklanmasıdır. Komşuların belirlenmesinde, örneklerin hepsi çok boyutlu uzaydaki vektörler olarak gösterilir. Genellikle Öklit mesafesi ölçütü kullanılsa da, Mahalanobis mesafesi gibi diğer ölçütler de kullanılabilir [85].

Verinin Gaussian dağılıma sahip olduğu varsayımında bulunan diğer sınıflandırıcıların aksine, KNN sınıflandırıcısı verinin sahip olduğu dağılıma bağımlı olmadan çalıştığından sınıflandırıcı olarak daha yüksek bir genellemeye imkan tanır [86]. Üstelik KNN algoritmasında eğitim verisinin kaybolmaması avantajına da sahiptir [87]. Fakat bu algoritmanın bazı dezavantajları bulunmaktadır. Öncelikle büyük veri setleri için sınıflandırmada test edilen veriyi tüm eğitim kümesiyle karşılaştırdığından oldukça uzun zaman harcamaktadır. Bu tıbbi teşhis uygulamaları gibi algoritmanın ürettiği doğru sonuçların ve çalışma süresinin önemli olduğu bazı alanlarda büyük sorun olacaktır [88]. Üstelik KNN algoritmasının başarımı gürültülü ve ilgisiz verilerin bulunduğu durumlarda ciddi bir şekilde etkilenmektedir [89]. Ayrıca verilerden elde edilen özniteliklerin ölçeklenmesi gerekmektedir [90].

2.7.2 Naive Bayesian Yöntemi

Naive Bayesian, hedef değişkenle bağımsız değişkenler arasındaki ilişkiyi analiz eden tahminci ve tanımlayıcı bir sınıflama algoritmasıdır [91]. Bayes karar kuralı, çeşitli kararlar arasındaki farkları, bu kararlara eşlik eden olasılıkları ve maliyetleri kullanarak ölçen istatistiksel bir yaklaşımdır [92].

Olasılık teorisi içinde incelenen bir olay olarak B olayına koşullu bir A olayı (yani B olayının bilindiği halde A olayı) için olasılık değeri, A olayına koşullu olarak B olayı (yani A olayı bilindiği haldeki B olayı) için olasılık değerinden farklıdır. Ancak bu iki birbirine ters koşulluluk arasında çok belirli bir ilişki vardır ve bu ilişkiye (ilk açıklayan istatistikçi İngiliz Thomas Bayes (1702–1761) dir [93].

Öznitelikler arasındaki bazı ilişkiler ve bağımlılıklar gösterilememiş olsa da Bayes karar verme kuralının birçok sınıflandırma probleminde oldukça başarılı sonuçlar verdiği görülmüştür [94].

Naive Bayes yaklaşımı, genellikle sonrasal olasılıkları hesaplamakta kullanılan ve iki rastgele olayın koşullu ve marjinal olasılıklarını ilişkilendiren bir teoremdir. Maksimum Olabilirlik ilkesi üzerine kurulu bir teoremdir. Bu durumda Bayes Teoremi, konulan olasılıkların doğruluk oranını hesaplamak için kullanılabilir.

A ve B rastgele olaylar olsun;

$$P(A \setminus B) = \frac{P(B \setminus A)P(A)}{P(B)} \quad (2—11)$$

P(A): A olayının bağımsız olasılığı

P(B): B olayının bağımsız olasılığı

P(B\A): A olayının olduğu bilindiğinde B olayının olasılığı (Likelihood, Şartlı Olasılık)

P(A\B): B olayının olduğu bilindiğinde A olayının olasılığı (Posterior, Artçıl Olasılık)

Naive Bayes kuralına dayanarak P(A\B) yi maksimum yapan durumlar hesaplanabilir.

‘ E ‘ A olayının bütün durumlarının kümesi;

$$A_{MAX} = \arg \max_{A \in E} P(A \setminus B) \quad (2—12)$$

$$A_{MAX} = \arg \max_{A \in E} \frac{P(B \setminus A)P(A)}{P(B)} \quad (2—13)$$

$$A_{MAX} = \arg \max_{A \in E} P(B \setminus A)P(A) \quad (2-14)$$

Burada P(B) sabit olarak göz ardı edilebilir. Sınıflandırıcının görevi yeni bir örnek için doğru sınıfı tahmin etmek olacaktır [95].

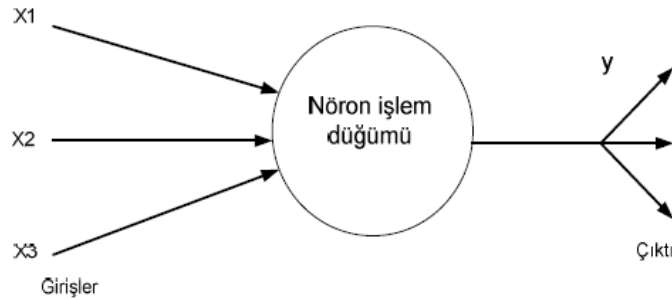
2.7.3 Yapay Sinir Ağları

Yapay sinir ağları, insan beyninin özelliklerini kullanarak, öğrenme yoluyla yeni bilgiler üretebilen ve keşfedebilen, önceki verilmiş örnekleri kullanarak yeni durumlara göre cevaplar üretebilen, kendi kendine karar verebilme yeteneği olan bilgisayar sistemleridir [96].

YSA, beyindeki sinirlerin varsayılan çalışma prensibini taklit ederek sistemlere öğrenme, genelleme yapma, hatırlama gibi yetenekler kazandırmayı amaçlayan bir bilgi işleme sistemidir. YSA, beynin bir işlevi yerine getirme yöntemini modellemek için tasarlanan bir sistem olarak tanımlanabilir. YSA, işlem elemanlarının birbirleri ile çeşitli şekillerde bağlanması ile oluşurlar. Sıklıkla katmanlı şekillerde tasarlanır. Beynin bilgi işleme yöntemine uygun olarak YSA, bir öğrenme sürecinden sonra bilgiyi toplama, işlem elemanları arasındaki bağlantı ağırlıkları ile bu bilgiyi saklama ve genelleme yeteneğine sahip paralel tasarlanmış bir işlemcidir [97].

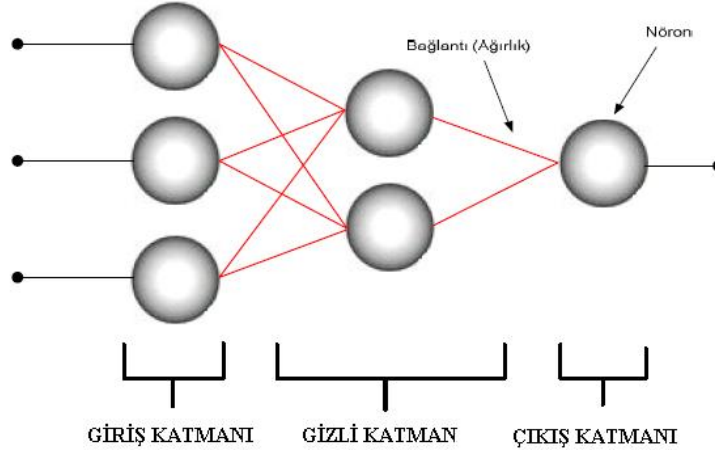
Yapay sinir ağı 3 bölümden oluşmaktadır. Bu bölümler sırasıyla nöron, bağlantılar ve ağırlıklardır.

Nöron: Biyolojik sinir ağlarında olduğu gibi yapay sinir ağlarında da temel etmen yapay nöron'dur (artificial neuron). Yapay nöron, bir ağın temel işlem elemanıdır. Ağ içinde yer alan tüm nöronlar bir veya birden fazla girdi alırlar ve tek bir çıktı verirler. Bu çıktı yapay sinir ağının dışına verilen çıktılar olabileceği gibi başka nöronlara girdi olarak da kullanılabilirler [98].



Şekil 2.16 Yapay nöron [98]

Bağlantılar: Yapay nöronları bir birine bağlayarak yapay sinir ağı oluştururlar. Aşağıdaki Şekil 2.17’de yapay nöron ve bağlantılar gösterilmiştir.



Şekil 2.17 Yapay sinir ağı modeli [98]

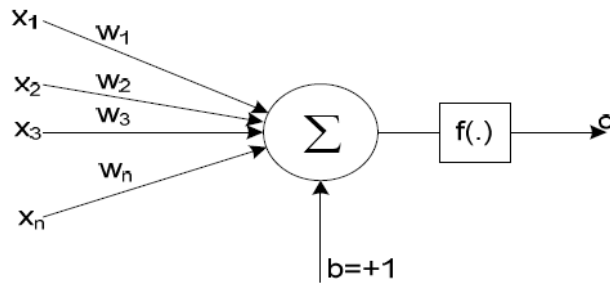
Nöronların aynı doğrultu üzerinde bir araya gelmeleriyle katmanlar oluşmaktadır. YSA modelin de üç adet katman bulunmaktadır. Birinci katman girdi katmanıdır. Giriş katmanının görevi dış ortamdan gelen verilerin yapay sinir ağına alınmasını sağlamaktır. İkinci katman ise çıktı katmanıdır. Bu katman da bilgilerin dışarıya iletilip aktarıldığı katmandır. Eğer girdi ile çıktı katmanları arasında katman varsa bu katmanlara gizli katman ismi verilir. Bir YSA’da gizli katman olması gerekliliği değildir veya birden fazla gizli katman da bulunabilir [99].

Ağırlıklar: Bir yapay sinir ağındaki en önemli unsurlardan biri de bağlantılardır. Nöronların bir birilerine veri aktarımını sağlayan bağlantılar aynı zamanda bir ağırlık değerine sahiptirler. Ağırlıklar bir nöronda girdi olarak kullanılacak değerlerin göreceli kuvvetini gösterir. Yapay sinir ağı içinde girdilerin nöronlar arasında iletimini sağlayan tüm bağlantıların farklı ağırlık değerleri bulunmaktadır. Böylelikle ağırlıklar her işlem elemanının her girdisi üzerinde etki yapmaktadır [100].

2.7.3.1 Yapay Sinir Ağının Çalışması

Bir yapay sinir hücresi biyolojik sinirlerde olduğu gibi, girişine gelen bilgileri değerlendirerek çıkışına verir. Bir yapay sinir hücresi, dış ortamdan veya diğer sinir hücrelerinden gelen bilgileri giriş olarak alır. Sinir hücresine gelen bilgiler bağlantı ağırlıkları ile ağırlıklandırılır ve her bir girişin hücre üzerindeki etkisi belirlenmiş olur. Ağırlıklandırma işlemi, giriş bilgi değerleriyle bağlantı ağırlık değerlerinin çarpımı

şeklinde elde edilir. Sinir hücresinin bilgiyi işleme ve yerine getirmesi biri toplam fonksiyonu diğeri geçiş fonksiyonu olmak üzere iki fonksiyon tarafından gerçekleştirilir. Hücreye gelen bu ağırlıklandırılmış bilgiler hücre içerisinde toplanarak birleştirilir ve hücreye gelen net bilgi elde edilir. Hücre içerisinde birleştirilen bu bilgiler daha sonra biyolojik sınırlarda olduğu gibi bir $f(.)$ fonksiyonu ile işlenir ve sonuçta yapay bir sinir hücresinin çıkışı elde edilir. Ayrıca, yapay sinir hücrelerine, net girdiyi artıran sabit +1 değerli bias girişi de uygulanabilir. Şekil 2.18'deki gibi elde edilen bir yapay sinir hücresinin matematiksel modeli aşağıdaki şekilde çıkarılabilir [98].



Şekil 2.18 Toplam ve aktivasyon fonksiyonlu bir yapay sinir ağı [98]

YSA'nın çıkışı aşağıdaki denklemle hesaplanır.

$$o = f(w \cdot x + b) \quad (2-15)$$

Buradaki, girişler x , ağırlar w matrisi ile toplu olarak gösterilir. n giriş sayısı olmak üzere;

$$x = x_1, x_2, x_3, \dots, x_n \quad (2-16)$$

$$w = w_1, w_2, w_3, \dots, w_n \quad (2-17)$$

Şeklinde yazılabilir. Bu değerler normalize edilirse;

$$net = \sum_{i=1}^n w_i x_i + b \quad ve \quad o = f(net) \quad (2-18)$$

$$o = f\left(\sum_{i=1}^n w_i x_i + b\right) \quad (2-19)$$

Şeklinde yazılabilir. Yukarıdaki formülde görülen $f(.)$ fonksiyonu, net girişlerin değerlendirildiği bir aktivasyon fonksiyonudur. Genelde doğrusal olmayan aktivasyon fonksiyonunun çeşitli tipleri vardır [98].

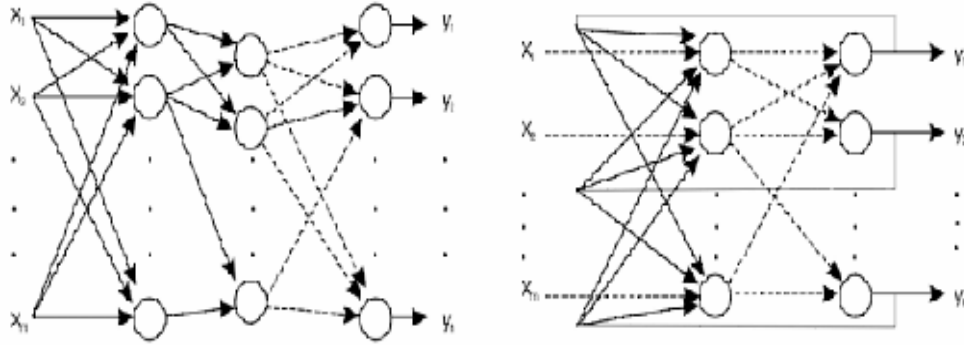
2.7.3.2 Radyal Temelli Fonksiyon Ağları

YSA'ların farklı çeşitleri mevcuttur. Bu farklılıkların sebebi kaynağı mimarisi, öğrenme yöntemi, bağlantı yapısı gibi özelliklerden kaynaklanmaktadır, YSA'lar üç temel kriter baz alınarak sınıflandırılırlar. Bu kriterlerden biri öğrenme yöntemine göredir. Temel olarak iki farklı öğrenme algoritması bulunmaktadır. Bu algoritmalar sırasıyla eğiticili ve eğitici-siz öğrenmedir [101].

Diğer bir sınıflandırma, ağıın kullandığı veriye göre yapılmaktadır. Temel olarak iki çeşittir. Bunlar Kalitatif verilerle çalışan ağlar, Kantitatif verilerle çalışan ağlardır.

Son sınıflandırma kriteri ise ağıın yapısına göredir. Ağıın yapısına göre sınıflandırma temel olarak iki çeşittir ve geri besleme yapısı içeren ve ileri besleme yapısı içeren ağlar olmak üzere ikiye ayrılmaktadır. İleri beslemeli sinir ağlarında işlem elemanları döngü oluşturmaz. Geri beslemeli sinir ağlarında işlem elemanları döngü içermektedirler. Döngü içerdikleri için ileri beslemeli sinir ağlarına oranla daha yavaş çalışırlar. Ayrıca her iki beslemeyi de içeren ağlar bulunmaktadır [102].

Şekil 2.19'da çok katmanlı ileri beslemeli ve çok tabakalı geri beslemeli ağ yapısı gösterilmiştir.



Şekil 2.19 Çok katmanlı ileri ve geri beslemeli ağların yapıları [101]

YSA çeşitleri arasında bilinirliği en yüksek olan ve kullanımı en çok olan ağlar geri yayılma ile eğitilen Geri-Yayılmalı Ağ (Backpropagation Network), Radyal Tabanlı Fonksiyon (Radial Basis Function), Hopfield ve Kohonen ağları gösterilebilir [101].

2.7.4 SOM Yapay Sinir Ağları: Özdüzenleyici Haritalar

Özdüzenleyici haritalar, yani SOM (Self-Organizing Maps), Fin profesör Tuevo Kohonen tarafından bulunan rekabet-temelli (competition) bir çeşit yapay sinir ağı modelidir [103]. Özdüzenleyici haritalar bazı kaynaklarda Kohonen Ağları olarak da anılmaktadır. Özdüzenleyici haritalar kümeleme işlemi ve gen ekspresyon örüntülerinin analizi için uygun bir tekniktir [104].

2.7.4.1 SOM Mimarisi

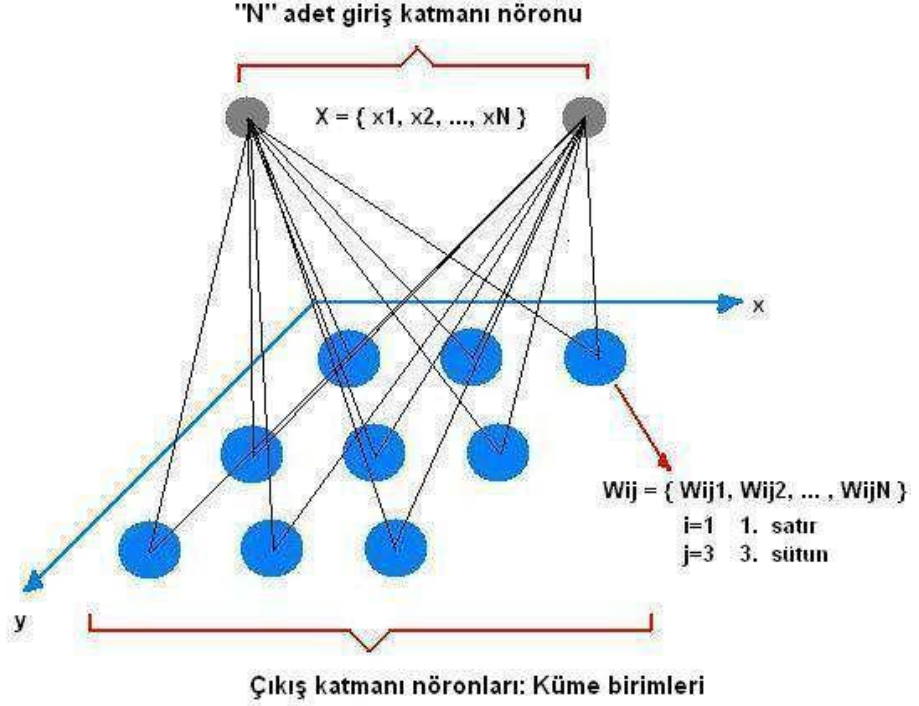
Özdüzenleyici haritalar, eğitici olmaksızın veri sınıflamayı öğrenebilen bir yapay sinir ağı modelidir. Bu nedenle gen kümeleme işlemi için son derece uygun bir alternatiftir. Çünkü gen kümeleme işlemi yapılırken, kullanılacak eğitim veri setleri bulunmamaktadır [105].

SOM sinir ağı iki temel bileşenden oluşmaktadır. Bunlar girdi birimleri ve küme merkezleridir (Şekil 2.20). SOM ağlarının girdisi, girdi verisinin öznitelikleridir (features) ve girdi birimlerinin adedi söz konusu özniteliklerin adedine eşittir. Rekabet eden nöronlar, kümeleme birimleridir. Kümeleme birimlerinin sayısı, diğer bir deyiş ile nöronların sayısı probleme ve kullanıcının seçimine göre değişmektedir. SOM ağlarında kümeleme birimleri arasında topolojik bir komşuluk ilişkisi olduğu varsayılmaktadır [105].

"N" adet özniteliğe sahip örnek veri

$X_i = \{ x_{i1}, x_{i2}, \dots, x_{iN} \}$, Örneğin: $x_1 = \{ 0.9, 0.3, \dots, 0.7 \}$

"N" adet özniteliğe sahip girdi verisi için tasarlanmış 3 x 3 SOM



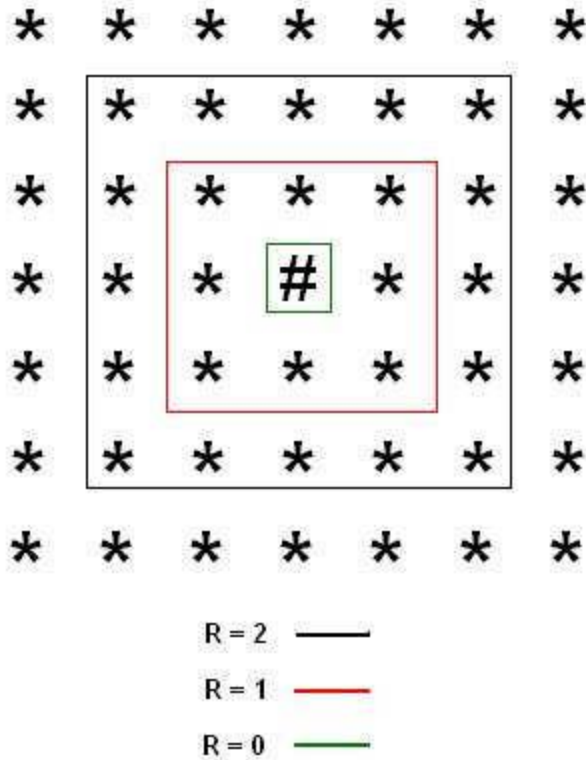
Şekil 2.20 N adet özniteliğe sahip girdi veri seti için tasarlanmış 3x3 boyutlu SOM'un yapısı [105]

Örnekte görülen 3x3 boyutlu SOM'da yer alan rekabet halindeki nöronların her biri, iki boyutlu uzayda belli bir pozisyona (x ve y koordinat bilgisine) sahiptir. Her nöronun bir ağırlık vektörü bulunmaktadır (W), ve bu vektör girdi vektörleri ile aynı boyuttadır. Örneğin, girdi verisi N boyutlu X vektörlerinden oluşuyor ise (x_1, x_2, \dots, x_N), her nöronun buna karşılık düşecek N boyutlu bir W ağırlık vektörü olacaktır (w_1, w_2, \dots, w_N) [105].

Nöron ağırlık vektörleri, temsil ettikleri küme ile ilişkilendirilmiş girdi verileri için bir model teşkil etmektedir. Özdüzenleme (self-organization) sürecinde, her girdi örneği için kendisine en çok benzeyen ağırlık vektörüne sahip nöron bulunur ve bu nöron kazanan nöron olarak seçilir. Bu seçim uzaklık fonksiyonuna göre yapılır. Bu nedenle uzaklık fonksiyonu seçimi kümeleme başarımını doğrudan etkiler. Kazanan

nörona ve kullanılan komşuluk topolojisi ile yarıçapına göre belirlenen komşu nöronlara ait ağırlık vektörleri güncellenir [106].

2, 1 ve 0 yarıçap değerleri ile dikdörtgensel topoloji için örnek komşuluk ilişkileri Şekil 2.21’de gösterildiği gibidir. Şekilde kazanan nöron ‘#’ sembolü ile gösterilirken diğer nöronlar ‘*’ sembolü ile gösterilmektedir. SOM sinir ağlarının komşuluk ilişkilerini belirtmek için dairesel, altıgen, vb. gibi farklı topolojiler de kullanılmaktadır [105].



Şekil 2.21 7x7 SOM üzerinde komşuluk ilişkilerinin dikdörtgensel topoloji ve 2, 1, 0 yarıçapları için gösterimi [105]

2.7.4.2 SOM Algoritması

SOM yapay sinir ağı modeline ait algoritma ve Kohonen öğrenme kuralı aşağıdaki gibidir [106].

1 boyutlu uzayda tanımlı SOM için, N adet özniteliğe sahip X girdi vektörleri $X_i=(X_{i1}, X_{i2}, \dots, X_{iN})$ olmak üzere ve W ağırlık vektörleri: $W_j=(W_{1j}, W_{2j}, \dots, W_{Nj})$ olmak üzere;

Adım 0: Ağırlıklara ilk değerlerini rastgele ata. Komşuluk parametrelerini (komşuluk topolojisi ve yarıçapını) ata. Öğrenim oranı (α) parametrelerini ata.

Adım 1: Sonlandırma koşulu sağlanmadığı sürece Adım 2-8'i gerçekleştir.

Adım 2: Her girdi vektörü X için Adım 3-5'i gerçekleştir.

Adım 3: Her j için, aşağıdaki hesaplamayı gerçekleştir: (Seçilen uzaklık fonksiyonunu kullanarak girdi verisi ile her nöron arasındaki mesafeyi hesapla. Bu örnekte Öklid uzaklığı kullanılmıştır.)

$$D(J) = \sum_{i=1}^N (w_{ij} - x_i)^2 \quad (2-20)$$

Adım 4: D(J) nin minimum olduğu J indeksini (yani kazanan nöronu) bul.

Adım 5: J'nin tanımlanan komşuluğu içerisinde kalan tüm j birimleri ve tüm i'ler için: $w_{ij}(\text{yeni}) = w_{ij}(\text{eski}) + \alpha[x_i - w_{ij}(\text{eski})]$. (Kohonen öğrenme kuralı)

Adım 6: Öğrenim oranını güncelle.

Adım 7: Belirlenen zamanlarda topolojik komşuluk yarıçapını azaltır. Adım 8: Sonlandırma koşulunu kontrol et. (İterasyon sayısına ulaşıldığında veya α eşik değere ulaştığında sona erdir.) [106].

Öğrenme oranı α , zamana veya devir sayısına göre yavaş yavaş azalan bir fonksiyondur. Laurene (1993)'de belirtildiği üzere öğrenme oranının lineer olarak azalan bir fonksiyon olması (örneğin geometrik azaltım gibi) pratik hesaplamalar için yeterlidir. Kazanan nörona komşu olan nöronların tespit edilmesi için kullanılan komşuluk yarıçapı değeri de kümeleme işleme süresince gerekli anlarda azaltılmalıdır [106].

3 GELİŞTİRİLEN ANALİZ YAZILIM MODÜLLERİ

3.1 Verilerin Elde Edilmesi

Çalışmadaki PSG kayıtları, Edirne Tıp Fakültesi Uyku Laboratuvarı'nda kullanılan 44 kanallı poligraf (Compumedics 44E serisi, Avustralya) ile tam PSG teknikleri kullanılarak elde edilen arşiv kayıtlarından alınmıştır. Bireye UPHB tanısı konulabilmesi için PLM (Periodic Leg Movement) indeksinin 15 ve üzeri olması gerekmektedir [107]. Çalışma farklı yaş ve cinsiyetten 2000 adet PSG kaydı içerisinden, UPHB teşhisi konulan 153 adet hastanın kayıtları kullanılarak gerçekleştirilmiştir.

PSG kaydında gece boyunca standart olarak kaydedilen EEG, EOG, EMG, EKG'nin yanında hava yolu, solunum eforu, oksijen saturasyonu gibi veriler de bulunmaktadır. PSG kayıtlarının uzunlukları yaklaşık olarak 7-8 saat arasında değişmektedir. Kayıtlar, PSG cihazına ait yazılımla otomatik olarak ya da uzman hekimlerin yaptığı görsel inceleme ile skorlanabilmektedir. Skorlanan bu kayıtlar SLP dosyası biçiminde saklanmaktadır. SLP dosya biçimi, EDF dosya formatına dönüştürülerek gerçekleştirilen yazılımda kullanılabilir hale getirilmiştir.

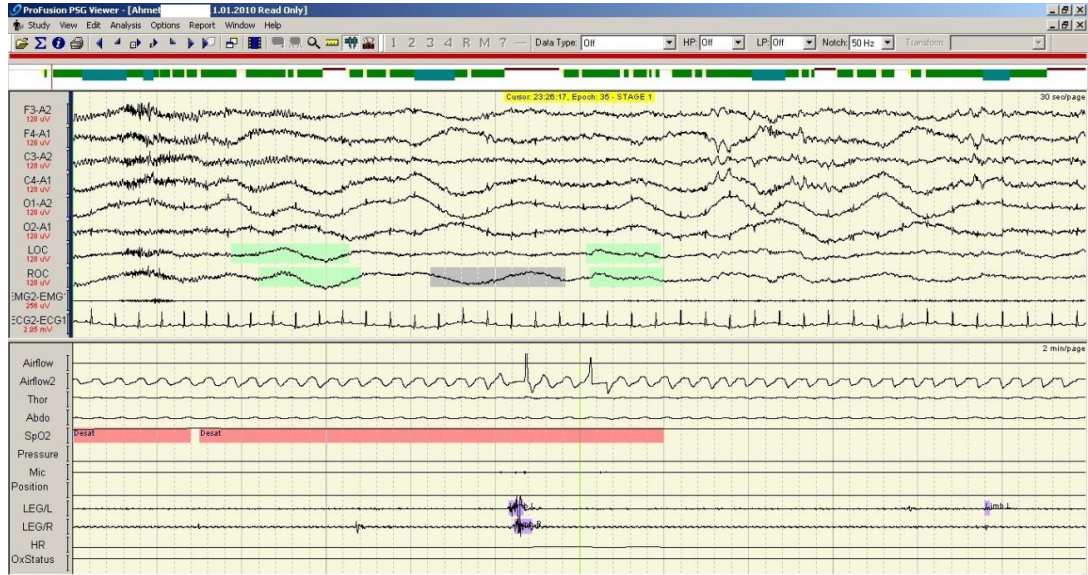
Avrupa Veri Biçimi (EDF) çok kanallı biyolojik ve fiziksel sinyallerin depolanması ve değiştirilmesi için basit ve esnek bir formattır. İlk olarak 1987 yılında Kopenhag'da yapılan uluslararası uyku kongresinde bir araya gelen medikal mühendisler tarafından geliştirilmiştir. 2002 yılında EDF formatının bir versiyonu EDF+ geliştirilmiştir. EDF+, EDF ile büyük ölçüde uyumludur. EDF+, EDF den farklı olarak kayıttaki ek açıklamaları, uyarıları ve olayları içermektedir [108]. EDF+'ta PSG verileri ve uyku skorlamaları bulunmaktadır. EDF sinir ileti çalışmaları, uyarılmış potansiyeller ve diğer veri toplama çalışmaları için kullanılabilir [109].

PSG montajında, 6 kanal EEG (F3-A2, F4-A1, O1-A2, O2-A1, C3-A2 ve C4-A1), sağ ve sol elektrookülografi (EOG) (LOC-A2 ve ROC-A1), çene elektromiyografisi (EMG), bacak elektromiyografisi (EMG) elektrokardiyografisi (EKG), pulse oksimetre ile parmak ucundan oksijen saturasyonu ölçümleri (SpO2), termistör

(üst solunum yolu sinyalleri için), toraks ve abdominal ölçümleri, horlama (mikrafon), vücut pozisyonu gibi bilgiler yer almaktadır. Bu bilgiler hastanın vücuduna uyku teknisyeni tarafından bağlanan (yerleştirilen) sensör ve elektrotlar yardımıyla alınmış ve kayıt edilmiştir.

EEG elektrotları uluslararası kabul görmüş olan 10–20 sistemine göre yerleştirilmiştir. EEG ve üst solunum sinyallerinin kaydedilmesi sırasında kullanılan örnekleme sayısı 256, akciğer ve karın solunum sinyallerinin örnekleme sayısı 128, yapılan filtrelemenin alt kesim frekansı 0,3 Hz, üst kesim frekansı ise 30 Hz olarak belirlenmiştir.

PSG verileri Amerikan Uyku Tıbbi Akademisinin (AASM) 2005 yılında yeniden düzenlediği uyku bozuklukları sınıflamasına göre ve Retschchaffen ve Kales rehberi kriterlerine göre Trakya Üniversitesi Tıp Fakültesi Öğretim Üyesi Prof. Dr. Levent Öztürk tarafından elle skorlanmıştır. PSG kaydı üzerindeki bütün skorlama zamanları, hangi olayın ne zaman başlayıp ne kadar sürdüğü gibi bilgiler XLM dosya formatında saklanmaktadır. Ham veriler ProFusionPSG programı kullanılarak XLM ve EDF dosya formatına dönüştürülmüştür. Aşağıda Şekil 3.1’de örnek bir PSG kaydı gösterilmiştir.



Şekil 3.1 PSG kaydı örneği

Çalışmada, farklı epok uzunluklarındaki verilerin sınıflandırmadaki etkisini test edebilmek için 2, 5, 60 saniyelik epoklar elde edilmiştir. PSG öznitelikleri ve üretilen sonuçlar olay bazında gerçekleştirilmiştir. Her bir PLM olayının başlangıç zaman

bilgisine karşılık gelen diğer kanallardaki PSG verilerinin kendine özgü öznelikleri epok şeklinde elde edilmiştir. Sınıflandırma için uyku dönemindeki PLM olanlar ile PLM olmayanlar arasında rastgele seçilen epoklar kullanılmıştır. Çalışmadaki epok süreleri, yazılım aracılığıyla 2, 5, 30, 60 saniye olarak seçilmiştir. Ayrıca apne esnasındaki ve apneden sonraki epoklar çalışmaya dahil edilmemiştir. Tablo 3.1’de 2, 5, 30, 60 saniyelik epok uzunları seçilerek bu epoklara ait, toplam epok sayısı, artefaktlı ve uyanık epok sayısı, artefaktsız ve uyku esnasındaki epok sayısı, PLM’nin bulunmadığı epok sayısı, PLM’nin bulunduğu epok sayısı bilgileri gösterilmiştir.

Tablo 3.1 Farklı epok sürelerine göre PLM’nin olduğu ve olmadığı, artefaktlı ve artefaktsız epoksayıları

Epok Süresi (sn)	Toplam Epok Sayısı	Artefaktlı ve Uyanık Epok Sayısı	Artefaktsız ve Uyku Esnasındaki Epok Sayısı	PLM’nin Bulunmadığı Epok Sayısı	PLM’nin Bulunduğu Epok Sayısı
2	1.921.815	521.927	1.399.888	1.346.609	53.279
5	768.726	185.319	583.407	543.704	39.703
60	64.018	14.687	49.331	37.230	12.101

3.2 Veritabanı Modülü

Veritabanı Modülü, geliştirilen yazılımın aynı zamanda anaformudur. Şekil 3.2’de modülün arayüzü gösterilmektedir. Şekil 3.2’de de görüldüğü gibi diğer modüllere bu modül üzerinden erişilir. Veritabanı Modülünde sırasıyla *EDF File*, *Database Name*, *Table Name*, *Signals*, *Pattern*, *Records* gibi bilgiler liste şeklinde görüntülenir. Veri tabanları ile ilgili işlemler için *Select File Path*, *Remove Database*, *Create Database*, *Create Analysis Table* butonları bulunmaktadır. Ayrıca diğer modüllere erişim için gerekli olan DSP Modül, XML Modül, Artifact Modül, Machine Learning Modül butonları bulunmaktadır.

Select File Path butonuna tıklanırsa; analiz için kullanılacak dosyaların bulunduğu klasör seçilir. Seçilen klasör içerisindeki ‘.edf’ uzantılı dosyaların isimleri sıralı bir şekilde EDF file liste kutusunda gösterilir.

Database Name altında, yaratılmış olan database adları bulunmaktadır. *Database Name* liste kutusundaki, veritabanı isimlerinde bulunan ‘_’ karakterinden sonraki rakam, epok süresini göstermektedir. Örneğin *psganalysis_2* adı ile oluşturulmuş olan veritabanındaki epok süresi 2 saniyedir. Bu liste kutusunda istenilen veritabanı seçilerek, aktif yapılabilir. Eğer aktif veritabanı değişirse, diğer listeler güncellenmektedir.

Table Name liste kutusunda ise *Database Name* liste kutusunda seçili, aktif olan veritabanındaki tabloların isimleri yer almaktadır. Burada da aynı şekilde istenilen tablo seçilerek, tablo içerisinde bulunan veriler görüntülenmektedir. (Şekil 3.2)

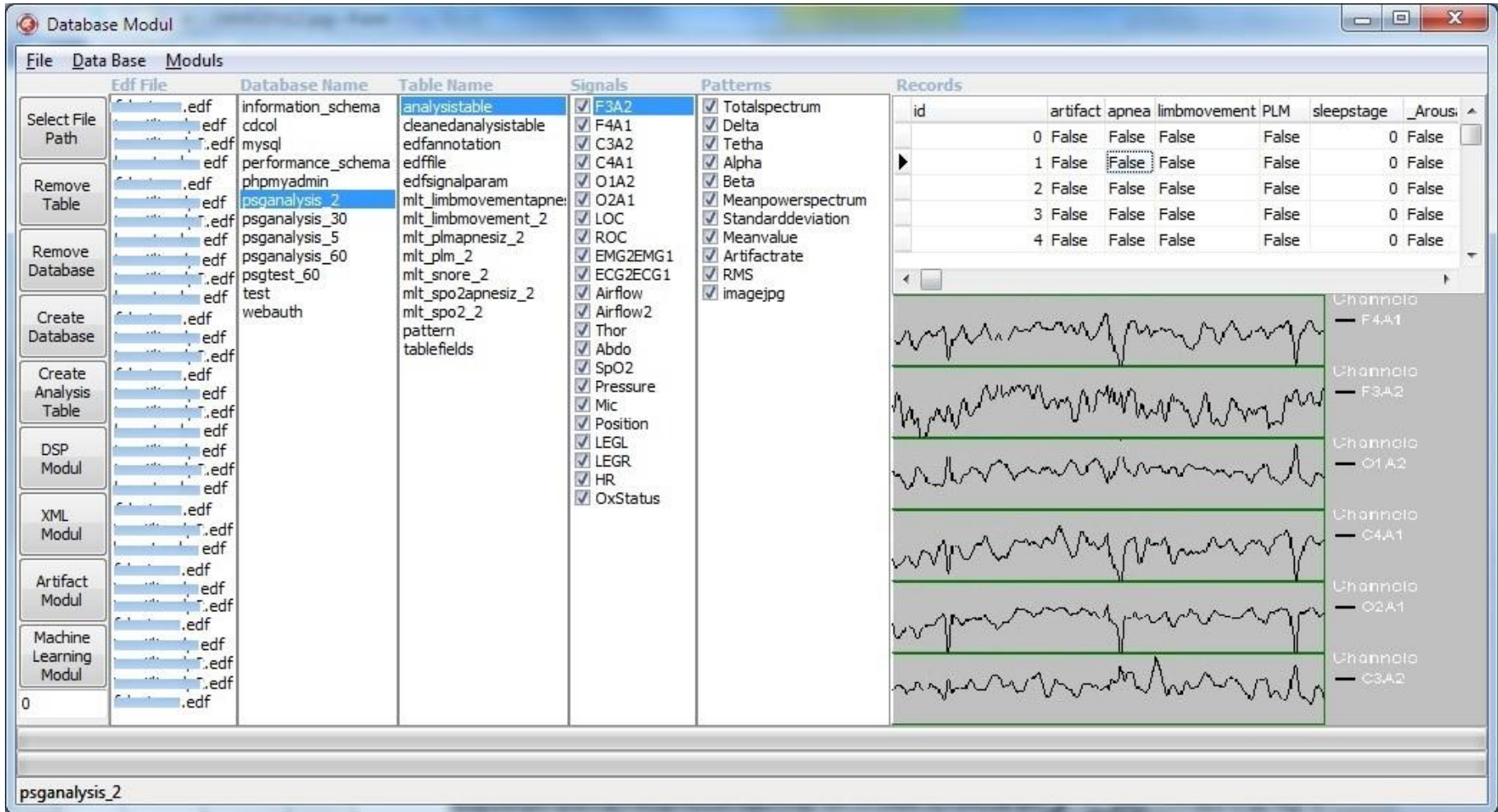
Signals seçenek liste kutusunda, seçili olan EDF dosyasında bulunan, sinyaller listelenmektedir. Burada istenilen sinyaller seçilerek analize dahil edilir ya da seçilmeyerek hariç tutulur. Burada bir sinyal seçildiğinde, *Pattern* seçenek liste kutusunda o sinyale ait öznitelikler listelenir. (Şekil 3.2)

Pattern seçenek liste kutusunda, analizde kullanılmak üzere istenilen öznitelikler seçilmektedir. (Şekil 3.2)

Records'ta ise table name liste kutusunda, seçili olan tablo verileri gösterilmektedir. *Records*'ta seçili epoga ait EEG sinyalleri, grafik olarak gösterilmektedir.(Şekil 3.2)

Modüldeki *Remove Database* butonuna tıklandığında, *Database Name* liste kutusunda seçili olan veritabanı silinmeden önce, bir uyarı mesajı gösterilir. Eğer silme işlemine devam etmek istenirse, aşağıdaki SQL komut satırı çalıştırılarak veritabanı silinir.

DROP DATABASE IF EXISTS: databasename;



Şekil 3.2 Veritabanı Modülünün ekran görüntüsü

3.2.1 Yeni Veritabanı Oluşturma

Oluşturulacak veritabanı ismi ve epok uzunluğu, kullanıcı tarafından giriş kutusuna girilir. Eğer aynı isimde bir veritabanı mevcutsa kullanıcıya uyarı mesajı olarak gösterilir. Veritabanı silinmek isteniyorsa aşağıdaki SQL komutu çalıştırılır.

```
DROP DATABASE IF EXISTS :databasename;
```

Daha sonra alttaki SQL komutu kullanılarak, yeni bir veritabanı ve analizde kullanılacak tablolar oluşturulur.

```
CREATE DATABASE IF NOT EXISTS :databasename;
```

```
USE :databasename;
```

```
CREATE TABLE IF NOT EXISTS `edfannotation` (  
  `fileno` int(11) DEFAULT NULL,  .  
  .  
  .
```

```
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

```
CREATE TABLE IF NOT EXISTS `edffile` (  
  `fileno` int(11) unsigned NOT NULL,  .  
  .  
  .
```

```
PRIMARY KEY (`fileno`)
```

```
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

```
CREATE TABLE IF NOT EXISTS `edfsignalparam` (  
  `signalno` tinyint(4) NOT NULL,  .  
  .  
  .
```

```
PRIMARY KEY (`signalno`)
```

```
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

```
CREATE TABLE IF NOT EXISTS `tablefields` (  
  `fieldalias` varchar(40) NOT NULL DEFAULT "",  
  .  
  .  
  .
```

```
PRIMARY KEY (`fieldalias`)
```

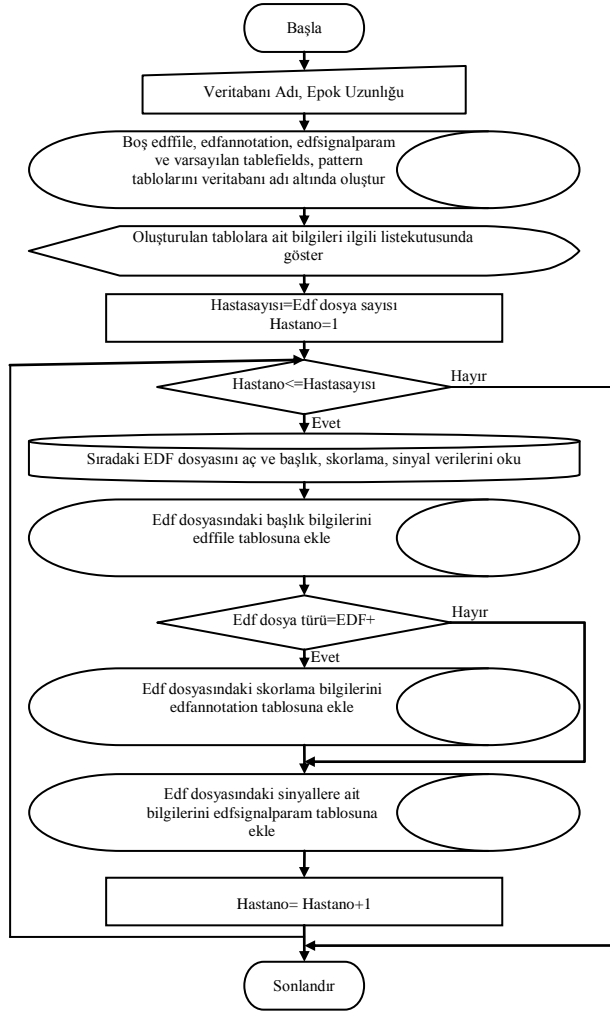
```
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

```
CREATE TABLE IF NOT EXISTS `pattern` (
  `patternno` int(10) unsigned NOT NULL,
```

```
PRIMARY KEY (`patternno`)
```

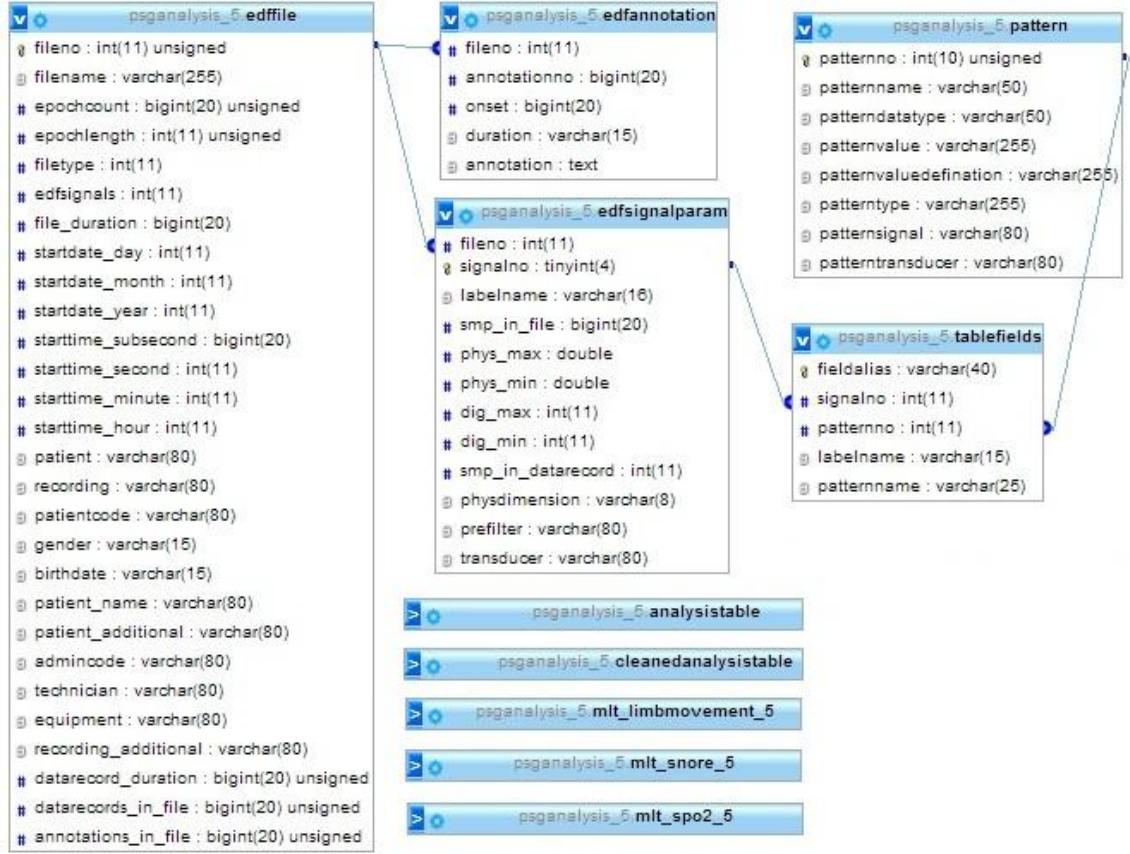
```
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

Şekil 3.3’de gösterilen akış şemasındaki işlemler gerçekleştirilerek, yapılacak analize özel yeni bir veritabanı oluşturulur.



Şekil 3.3 Yeni veritabanı oluşturma akış şeması

Yapılan işlemlerin sonunda, epok süresi 5 saniye olan örnek bir analiz ve veritabanına ait tablolar ve arasındaki ilişkiler Şekil 3.4’de gösterilmektedir. Şekil 3.4’de diğer modüller tarafından oluşturulan tablolar, çok fazla alan içerdiği için kapalı olarak gösterilmiştir.



Şekil 3.4 Epok süresi 5 sn olan örnek bir analiz ve veritabanında bulunan, tablolar ve arasındaki ilişkiler

3.2.2 Analiz Tablosu Oluşturma

Seçili olan veritabanındaki EDF dosyalarının türüne ve içeriğine göre seçenek liste kutusu 1’de sinyal isimleri listelenir. Seçenek liste kutusu 2’de ise seçenek liste kutusu 1’deki seçili olan sinyale ait öznitelik isimleri listelenir. Analizde kullanılacak sinyaller seçenek kutusu 1’den, parametreler ise seçenek kutusu 2’den seçilmektedir. Her seçimde seçilen sinyal ve özneliğe ait bilgiler, aşağıdaki örnek SQL komutu kullanılarak ‘tablefields’ tablosuna eklenir.

```
INSERT INTO `tablefields` (`fieldalias`, `signalno`, `patternno`, `labelname`,  
`patternname`) VALUES ('C3A2_RMS', 0, 9, 'C3A2', 'RMS');
```

Seçim kaldırıldığında ise alttaki SQL komutu çalıştırılır.

```
'DELETE FROM `tablefields` WHERE `signalno` =' + signalno + ' AND  
`patternno` =' + patternno;
```

Yazılıma bacak hareketi olaylarının yanı sıra, PSG verileri ile farklı analizlerin yapılabileceği modüler bir yapı kazandırılmıştır. Bütün seçimler tamamlandığında *Create Analysis Table* butonuna tıklanarak, DSP ve XML Modüllerine eklenecek olan kayıtların alan isimleri ve değişken türleri, özniteliğin türüne göre SQL komutları kullanılarak oluşturulur.

```
DROP TABLE IF EXISTS `analysistable` ;CREATE TABLE `analysistable` (  
  `id` bigint unsigned NOT NULL ,  
  `artifact` boolean NOT NULL ,  
  `apnea` boolean NOT NULL ,  
  `limbmovement` tinyint NOT NULL ,  
  `PLM` boolean NOT NULL ,  
  `sleepstage` tinyint unsigned NOT NULL default 0 ,  
  `_Arousal` boolean NOT NULL default 0 ,
```

Tablefields tablosundaki kayıt sayısı kadar, sırasıyla *tablefields*. *fieldalias`* veritürü ise pattern tablosundan *pattern.patnerdatatype* alanlarındaki veriler üstteki SQL satırlarının altına eklenir.

```
tablefields. fieldalias` pattern.patnerdatatype,
```

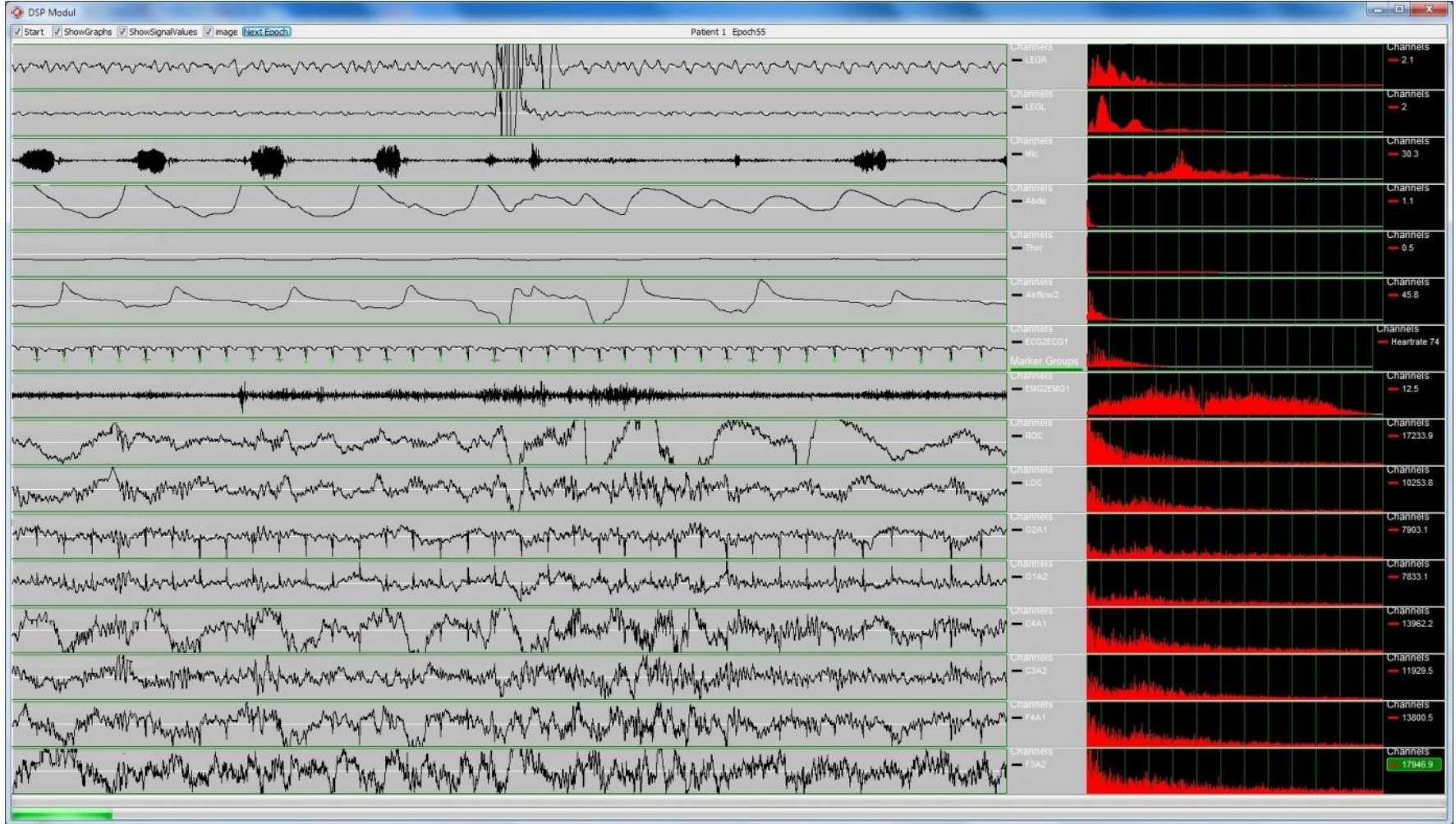
Komutun son satırına alttaki SQL komutu eklenerek çalıştırıldığında, analiz tablosu oluşturulur.

```
,PRIMARY KEY (`id`)) ENGINE=innodb DEFAULT CHARSET=utf8;');
```

3.3 Dijital Sinyal İşleme Modülü

Çalışmada EEG sinyallerinin alt frekans bantlarının (Delta, Teta, Alfa, Beta) elde edilmesinde Haar dalgacık dönüşüm tekniği kullanılmıştır. Elde edilen alt frekans bantlarının frekans spectrum değerlerini elde etmek için Ayırık Zamanlı Fourier Dönüşüm (Discrete Fourier Transform) tekniği kullanılmıştır. Bu teknik ile zaman düzlemindeki EEG sinyalleri frekans düzlemine dönüştürülmüştür. Bu yöntem ayrıca kalp atım oranı (Heart rate) ve horlamaların tespit edilmesinde de kullanılmıştır.

DSP Modülünün arayüzü Şekil 3.5’de gösterilmektedir. Bu modülün sol tarafında analize dahil edilen biyoelektrik sinyaller, sağ tarafında ise spektrum analizi ya da sinyale özel değerler grafik olarak gösterilmektedir. Bu modülün sol üst köşesinde *Start*, *ShowGraphs*, *ShowSignalValues*, *Image* gibi kontrol kutuları aktif ve pasif yapılabilmektedir. *Next Epoch* butonuna tıklanarak, gerçekleşen işlemler adım adım takip edilebilmektedir. Modül ekranın üst orta kısmında ise hasta numarası ve bu hastanın hangi epogunun işlem gördüğü, işlem barında ve yazılı olarak gösterilir.

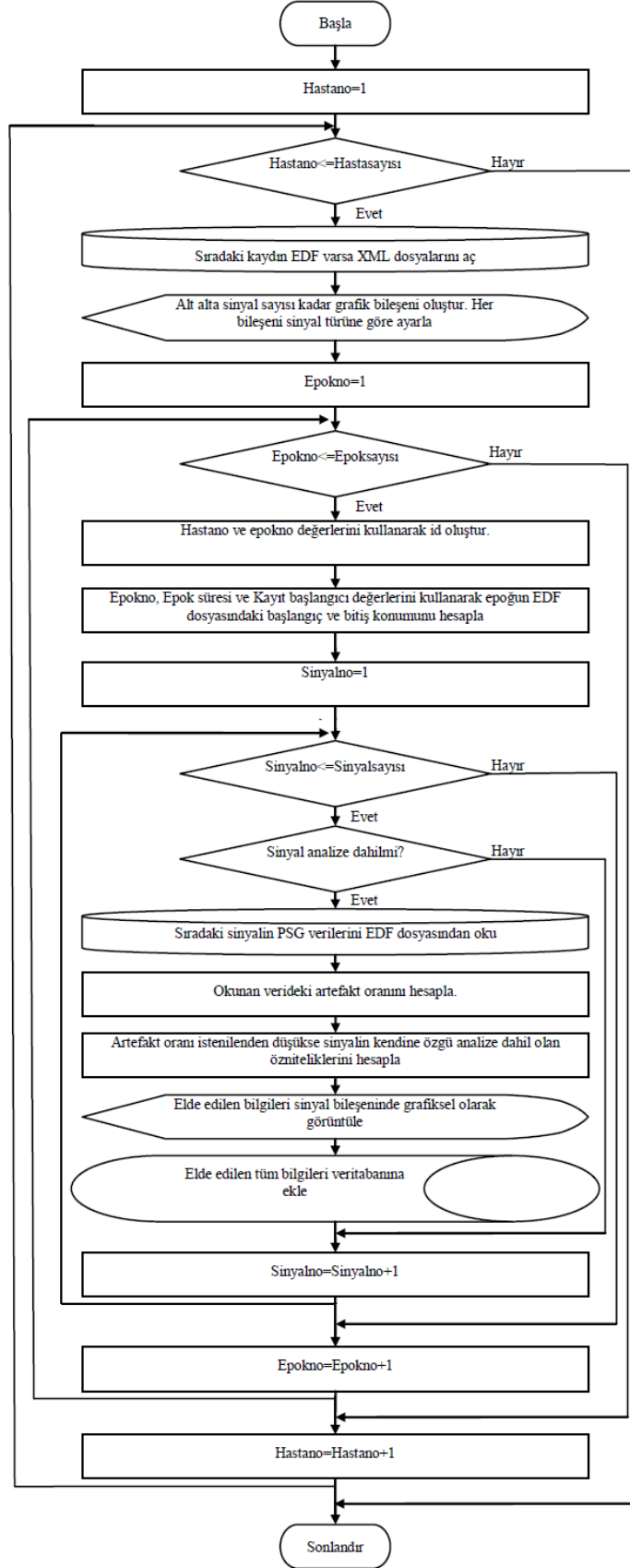


Şekil 3.5 DSP Modülünün ekran görüntüsü

Analizde kullanılmak üzere, Veritabanı Modülünde seçilen parametreler elde edilerek *analysistable* tablosuna *id* değeriyle yazılır. *Id* değeri, hasta ve epok numarası ile elde edilir. Modülün genel akış çizelgesi Şekil 3.6’da anlatılmaktadır.

Modülde kullanılan bileşenler, dinamik olarak çalışma esnasında oluşturulmaktadır. Yani yapılan analizde, kullanılacak parametre ve sinyallerin türüne göre değişmektedir.

Modülün çalışma hızı; hasta sayısına, epok süresine, sinyal ve parametre sayısına göre değişmektedir. Modülün daha hızlı çalışması isteniyorsa *Show Graphs* ve *Show signal value* seçenek kutuları pasif duruma getirilerek, grafiklerin görüntülenmesi engellenebilmektedir.



Şekil 3.6 DSP Modülünün akış şeması

3.3.1 Öznitelikler

Bu modülde analize dahil edilen sinyallerin kendine özgü öznitelikleri, DSP ve istatistiksel yöntemler kullanılarak elde edilerek, *analysisstable* tablosuna eklenmektedir. Tablo 3.2’de yazılımda kullanılan tüm parametrelerin özellikleri verilmektedir.

Tablo 3.2 Modülde kullanılan öznitelikler ve değerleri (kesirli sayı:1, tamsayı: 2)

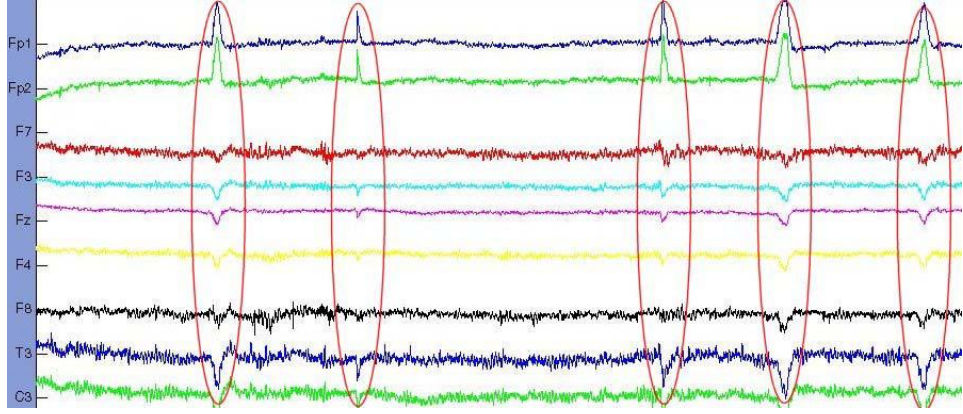
No	Adı	Veri Tipi	Değer	Değer Tanımı	Tipi	Sinyali	Çevirici
0	Güç Spektrumu	1			Spectral		EEG
1	Delta	1	0-100	%	Spectral		EEG
2	Teta	1	0-100	%	Spectral		EEG
3	Alfa	1	0-100	%	Spectral		EEG
4	Beta	1	0-100	%	Spectral		EEG
5	Ortalama Güç	1			Spectral		EEG
6	Standart Sapma	1			İstatistiksel	Tümü	
7	Ortalama Değer	1			İstatistiksel	Tümü	
8	Artefakt Oranı	1	0-100	%	Artifact	Tümü	
9	RMS	1			İstatistiksel	Tümü	
10	Yorulma Oranı	1			İstatistiksel		EMG
11	Değer	1			Ham Veri		Oksimetre
12	Apne	2	0,1	Var,Yok	Değerlendirilmiş, İstatistikse	Hava Akımı2	
13	Obstructive apne	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
14	Hypopne	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
15	MixedApne	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
16	CentralApne	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
17	ObstructiveHypopne	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
18	MixedHypopnea	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
19	CentralHypopnea	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
20	RERA	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
21	Unsure	2	0,1	Var,Yok	Değerlendirilmiş	Hava Akımı2	
22	Solumun artefaktı	2	0,1	Var,Yok	Değerlendirilmiş, Artefakt	Hava Akımı2	
23	Sol Bacak Hareketi	2	0,1	Var,Yok	Değerlendirilmiş	LEGL	
24	Periyodik Bacak	2	0,1	Var,Yok	Değerlendirilmiş	LEGL	
25	Sağ Bacak Hareketi	2	0,1	Var,Yok	Değerlendirilmiş	LEGR	
26	Periyodik Bacak	2	0,1	Var,Yok	Değerlendirilmiş	LEGR	
27	Oksijen	2	0,1	Var,Yok	Değerlendirilmiş	SpO2	
28	Oksijen Satürasyonu	2	0,1	Var,Yok	Değerlendirilmiş	SpO2	
29	Bradycardia	2	0,1	Var,Yok	Değerlendirilmiş	ECG1	
30	Tachycardia	2	0,1	Var,Yok	Değerlendirilmiş	ECG1	
31	Solumun Paradoksu	2	0,1	Var,Yok	Değerlendirilmiş	Thor	
32	Uyanma	2	0,1	Var,Yok	Değerlendirilmiş		
33	TcCO2 Artefaktı	2	0,1	Var,Yok	Değerlendirilmiş, Artefakt		
34	EtCO2	2	0,1	Var,Yok	Değerlendirilmiş, Artefakt		
35	DistalpH	2	0,1	Var,Yok	Değerlendirilmiş, Artefakt		
36	DistalpH	2	0,1	Var,Yok	Değerlendirilmiş		
37	ProksimalpH	2	0,1	Var,Yok	Değerlendirilmiş, Artefakt		
38	Kan Basıncı Artefaktı	2	0,1	Var,Yok	Değerlendirilmiş, Artefakt		
39	Vücut Sıcaklığı	2	0,1	Var,Yok	Değerlendirilmiş, Artefakt		
40	En Düşük SpO2	2	0-100		İstatistiksel, İstatistiksel	Spo2	
41	Desatürasyon	2	0-100		İstatistiksel, İstatistiksel	Spo2	
42	Pozisyon	2	0,1,2,	Ön, Arka, Sol, Sağ	Ham Veri	Pozisyon	
43	Oksijen Durumu	2	0,1,2,	İyi,Orta,Zayıf,Çok	Ham Veri		Oximeter
44	Işık	2	0,1	Kaplı,Açık	Ham Veri		
45	İnsan pozisyonu	2	0,1,2,	Ön, Arka, Sol, Sağ	Ham Veri	Pozisyon	
46	Uyku Evreleri	2	0,1,2,		Değerlendirilmiş		EEG
47	Horlama	2	0,1,2		Spectral		Sound
48	Periyodik Nefes	2	0,1		Değerlendirilmiş	Abdo	
49	Nabız	1			İstatistiksel		Oximeter
50	Artefakt	1	0,1	Var,Yok	Artefakt	Tümü	
51	imagejpg	Blob		jpg		Resim	All
52	imagebmp	Blob	bmp		Resim	Tümü	

Modüle yeni bir öznitelik eklemek istendiğinde, özniteliğin *patternno* (öznitelik numarası), *patternname* (öznitelik adı), *patterndatatype* (özniteliğin SQL veritabanındaki veri tipi), *patternvalue* (özniteliğin varsa sayısal veri aralığı), *patternvaluedefination* (sayısal veri aralığının sözlü karşılığı), *patternrype* (öznitelik belli bir gruba aitse grup adı), *patternsignal* (özniteliğin kullanılabileceği sinyaller), *patterntransducer* (özniteliğin kullanılabileceği transdüserler) özellikleri *pattern* tablosuna eklenir. Herhangi bir yazılım geliştirme ortamıyla geliştirilen, özniteliğin adıyla oluşturulan DLL dosyası, modüle özel giriş ve çıkış parametreleri kullanılarak modüle dahil edilebilir. Bu özelliği sayesinde yazılım, tek bir yazılım geliştirme ortamına bağımlı kalmaz. Gelecekte farklı kişi ve yazılım geliştirme ortamları kullanılarak geliştirilebilir.

3.4 Artefakt Modülü

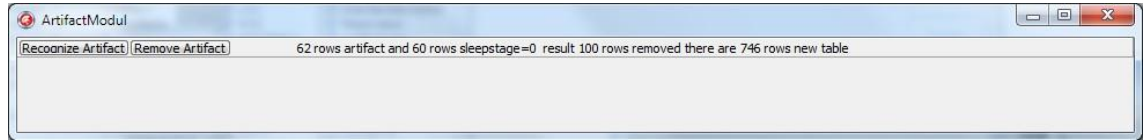
Elektroensefalografi işaretleri, beyin yüzeyinden elektrotlarla algılanan düşük genlikli biyoelektrik işaretlerdir. Bu işaretlerin genliği tepeden tepeye 2-100 mikrovolt ve frekans bandı ise 0.1-60 Hz aralığında bulunmaktadır. Ama EEG işaretleri düzensiz bir yapıya sahip olmasından dolayı daima artefakt bulundurlar. Genelde artefaktlar sinyaldeki istenmeyen verilerdir. Bu artefaktlar hareketli uzuvların örneğin; gözün hareketlerinden oluşan ya da vücudun dış etkilere gösterdiği tepki sonucu oluşan biyoelektrik potansiyellerdir. EEG’de oluşan Elektrookülografi (EOG) ve Elektromyografi (EMG) artefaktları, fizyolojik artefaktların en önemli kaynaklardan biri olarak kabul edilir [110].

EEG sinyalleri düzensiz bir yapıya sahip olduğundan bünyelerinde daima artefakt bulundurlar. Artefakt sinyaldeki istenmeyen bozucu etkilere verilen isimdir. Bu artefaktlar hareketli uzuvların, gözün, terlemenin oluşturduğu ya da vücudun dış etkilere gösterdiği tepki sonucu oluşan biyoelektrik potansiyellerdir. EEG verilerini etkileyen tüm bu olumsuzlukların giderilmesi için önce onları tanımak ve tespit etmek gerekir [111].



Şekil 3.7. Göz kırpması sonucu EEG dalgasında oluşan artefakt [111]

Şekil 3.7’de göz kırpması nedeniyle EEG dalgasında oluşan artefaktlar görülmektedir. Kaydedilen sinyallerin belirli bir kısmında artefaktlar gözlemlendiğinde, sinyalin hasarlı kısmını çıkarıp atmak bir çözüm olabilmektedir fakat bu esnada sinyal içerisinde faydalı bilgiler de atılmış olabilir. Ayrıca araştırma amaçlı EEG verileri çok kısa aralıklarla kaydedildiğinden sinyalin kesilmesi büyük veri kayıplarına yol açabilmektedir. Bu nedenle sinyalin kesilip atılması yerine sinyal bünyesinden artefaktları gidermek için filtreler veya yüksek mertebeden istatistiksel algoritmalar denenmesi daha etkili bir yoldur [111].



Şekil 3.8 Artefakt Modülünün arayüzü

Modülün arayüzü Şekil 3.8’de görüldüğü gibi görsel olarak tasarlanmamıştır. Bilinen artefakt tespit etme yöntemleri kullanılarak, PSG kaydında bulunan artefaktlar tespit edilmektedir. Epoktaki artefakt yüzdesi, istenilen değerden düşük ise bu epokun kaydı *cleanedanalysistable* tablosuna eklenir.

Analize hastanın sadece uyku veya uyanıklık dönemlerindeki veriler dahil edilmek istenirse, bu modülde *cleanedanalysistable* tablosundaki, uyku veya uyanıklık dönemlerindeki epoklar isteğe bağlı olarak tabloya eklenebilir veya tablodan çıkartılabilir. Tüm bu işlemlerin istatistiksel sonuçları form arayüzünde gösterilmektedir.

3.5 XML Modülü

Dokümanların organizasyonu için, 1960'lı yıllarda IBM tarafından önerilen GML (Generalize Markup Language) işaretleme dili temel alınarak SGML (Standard Generalized Markup Language) işaretleme dili geliştirilmiş ve bu uzun süre standart doküman tanımla dili olarak kullanılmıştır [112]. XML dili de W3C (World Wide Web Consortium) tarafından, SGML diline uygun bir işaretleme dili olarak ortaya çıkmıştır [113].

Günümüzde verilerin saklanması için XML (Extensible Markup Language) dokümanlarının kullanılması konusunda pek çok çalışma yapılmakta hatta bazı ilişkisel veritabanları XML dokümanlarını yönetebilir hale gelmektedir. XML dokümanlarının saklanması ve XML dokümanları üzerinde sorgulamaların yapılabilmesi için yapılan çalışmalar sonuçlanmaya başlamış ve ilişkisel veritabanları ile yarışabilmek için ilk adım atılmıştır [114].

Aşağıdaki şekilde örnek bir XML dokümanı verilmiş ve genel özellikleri anlatılmıştır [114].

```

<?xml version="1.0" encoding="UTF-8"?>
- <CMPStudyConfig>
  <EpochLength>30</EpochLength>
  - <StepChannels>
    - <StepChannel>
      <Input>Position</Input>
      - <Labels>
        <Label>Front</Label>
        <Label>Back</Label>
        <Label>Left</Label>
        <Label>Right</Label>
      </Labels>
    </StepChannel>
    :
  </StepChannels>
  - <ScoredEventSettings>
    - <ScoredEventSetting>
      <Name>PLM (Left)</Name>
      <Colour>0</Colour>
      <TextColour>0</TextColour>
      <Input>LEG/L</Input>
    </ScoredEventSetting>
    - <ScoredEventSetting>
      <Name>PLM (Right)</Name>
      <Colour>0</Colour>
      <TextColour>0</TextColour>
      <Input>LEG/R</Input>
    </ScoredEventSetting>
    + <ScoredEventSetting>
      :
    </ScoredEventSetting>
  </ScoredEventSettings>
  - <ScoredEvents>
    - <ScoredEvent>
      <Name>PLM (Right)</Name>
      <Start>190</Start>
      <Duration>4.1</Duration>
      <Input>LEG/R</Input>
    </ScoredEvent>
    - <ScoredEvent>
      <Name>PLM (Left)</Name>
      <Start>267.9</Start>
      <Duration>0.6</Duration>
      <Input>LEG/L</Input>
    </ScoredEvent>
    :
  </ScoredEvents>
  - <SleepStages>
    <SleepStage>0</SleepStage>
    <SleepStage>2</SleepStage>
    :
  </SleepStages>
</CMPStudyConfig>

```

Şekil 3.9 Örnek bir XML dokümanı

XML dokümanının sürüm ve kodlama bilgileri için dokümanın en üstünde `<?xml version="1.0" encoding="utf-8" ?>` etiketi tanımlanır. Etiketler `<ve>` karakterleri ile çevrelenir.

Örnekte *CMPStudyConfig* adında kök etiket tanımlanmış ve bu kök etiket içinde *EpochLenght*, *StepChannels*, *ScoredEventSettings*, *SleepStages* etiketleri eklenmiştir. Veriler bu etiketler arasında bulunmaktadır. Verilerin hiyerarşik bir şekilde tanımlanması gerektiği için her bir verinin bir kök etiketi bulunması ve diğer etiketlerin

kök etiketin altında veya daha alt etiketlerinde tanımlanması gerekmektedir. Bir etiket arasında, başka etiketlerden istenildiği kadar tanımlanabilmektedir. Tanımlanan her etiketin verisi girildikten sonra kapatılması gerekmektedir [114].

Yukarıda Şekil 3.9’da gösterildiği gibi *ScoredEventSettings* etiketi altında *ScoredEventSetting* etiketi tanımlanmıştır. *ScoredEventSetting* etiketi altında da *Name*, *Start*, *Duration*, *Input* etiketleri tanımlanmıştır. Bu etikeler sırasıyla olay adı ‘PLM(Right)’, olayın hangi saniyede başladığı ‘190’, olayın ne kadar sürdüğü ‘4.1’, girişin ne olduğu ‘LEG/R’ olarak gösterilmiştir. İşaretleyici olan veya verisi bulunmayan etiketler aynı etiket içinde kapatılabilmektedir. Örnekte *ScoredEventSettings* etiketi veri içermediği için `< ScoredEventSettings />` olarak tanımlanmıştır. Burada etiket tanımlanmış ve aynı etiket içinde kapatılmıştır. Bu gösterim `< ScoredEventSettings > </ ScoredEventSettings >` ile aynı anlama gelebilmektedir [114].

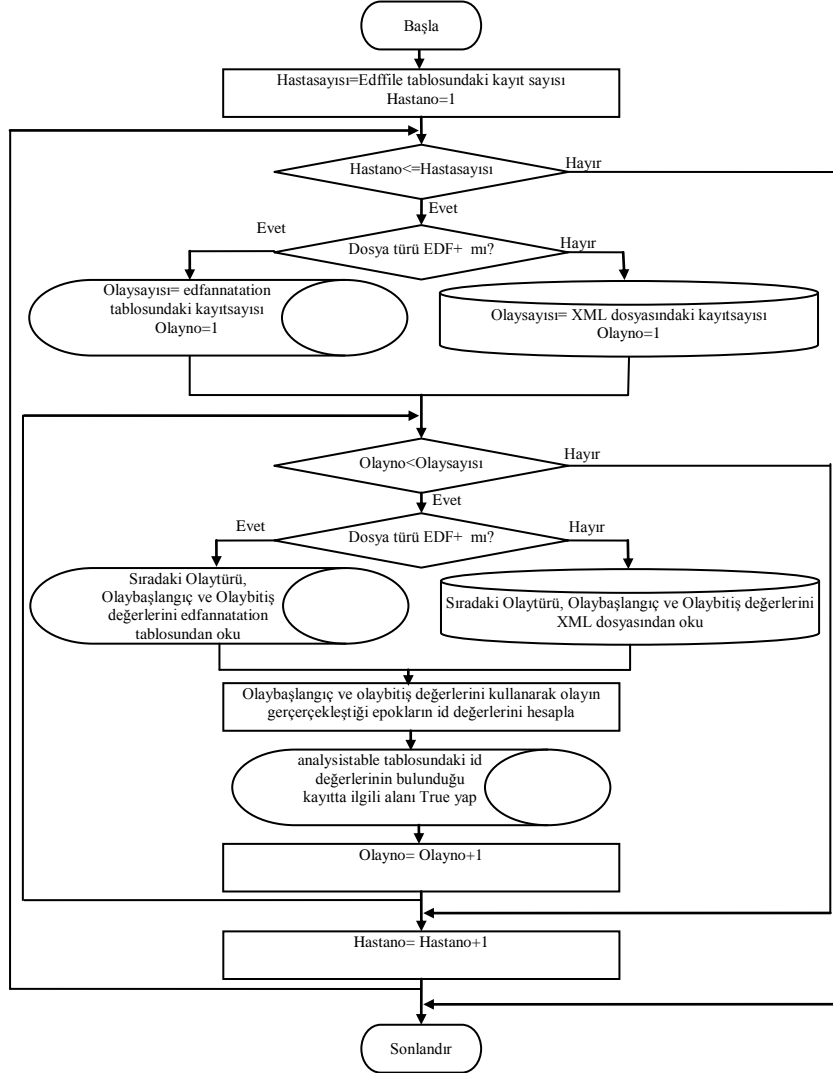
XML Modülünün arayüzü Şekil 3.10’deki gibidir. ProFusionPSG programı ile hasta kayıtları ‘.edf’ formatına dönüştürüldüğünde, EDF formatlı dosyanın yanı sıra XML formatında bir dosya da oluşturulmaktadır. Bu XML dosyasında uzman hekim tarafından skorlanan, hastaya ait Position, OxStatus, Central Apnea, Obstructive Hypopnea, Hypopnea, Mixed Hypopnea, RERA, SpO2 desaturation, Respiratory artifact, SpO2 desaturation, Arousal, PLM, Limb Movement gibi olaylar ve uyku evresi bilgileri bulunmaktadır. PSG kaydı üzerindeki bütün skorlama zamanları, hangi olayın ne zaman başlayıp ne kadar sürdüğü gibi bilgiler XLM dosya formatında saklanmaktadır. Analiz edilen epok süresine denk gelen herhangi bir olay varsa o anki epok parçasına denk gelen olay bilgileri XML dosyasından elde edilmektedir.



Şekil 3.10 XML Modülünün arayüzü

Bu olayların isimleri, başlangıç ve olay süreleri XML dosyasının içinde bulunmaktadır. Eğer oluşturulan EDF dosyasının türü EDF+ ise XML dosyasındaki

bilgilerde EDF dosyası içerisinde *annotation* olarak saklanmaktadır. *Analysis* tablosunda, okunan olayın gerçekleştiği zamandaki epok kayıtları güncellenir. Modülün akış şeması Şekil 3.11’de ayrıntılı olarak anlatılmaktadır.



Şekil 3.11 XML (Skorlama) Modülü arayüzü

3.6 Makine Öğrenmesi Modülü

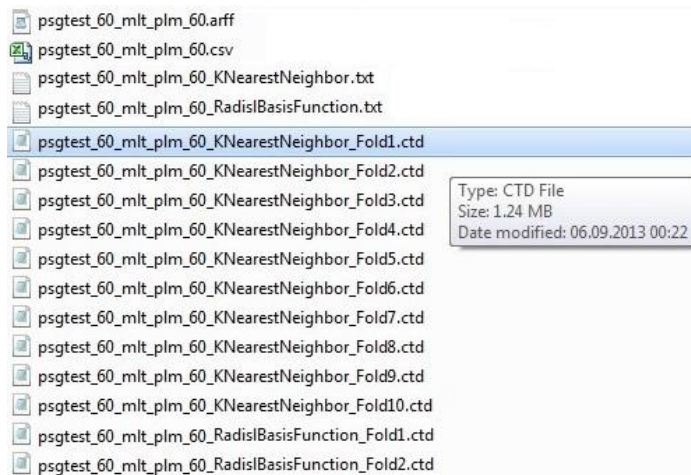
Makine öğrenmesi modülündeki amaç, beyinden ‘F3-A2’, ‘F4-A1’, ‘C3-A2’, ‘C4-A1’, gözlerden ‘LOC’, ‘ROC’, çeneden ‘EMG2-EMG1’, kalpten ‘ECG2-ECG1’, solunum organlarından ‘Airflow1’, ‘Airflow2’, ‘Thor’, ‘Abdo’, kandaki ‘Spo2’,

horlama sesinden ‘Mic’, yatış pozisyonundan ‘Position’, ‘OxStatus’ ve ‘LEG’ gibi sinyallerden alınan özniteliklerin, birkaçını veya hepsini makine öğrenmesine giriş, sadece bir özniteliği çıkış parametresi olarak vererek, girişteki özniteliklerin çıkıştaki öznitelikleri ile ilişkisini ortaya koymaktır.

Yapılan literatür araştırmaları sonucunda, modüle en çok kullanılan ve bilinen makine öğrenmesi algoritmaları eklenmiştir. Bu modül de modüler yapıda tasarlanmıştır. Modüle yeni bir algoritma eklenecek ise algoritmanın DLL’si ve parametreleri tanımlanarak dahil edilebilir. Ayrıca veriler, başka bir yazılım tarafından çalıştırılmak üzere ARFF ve CSV gibi dosya biçimlerine aktarılabilir ya da bu dosya biçimlerinden veritabanına alınabilir.

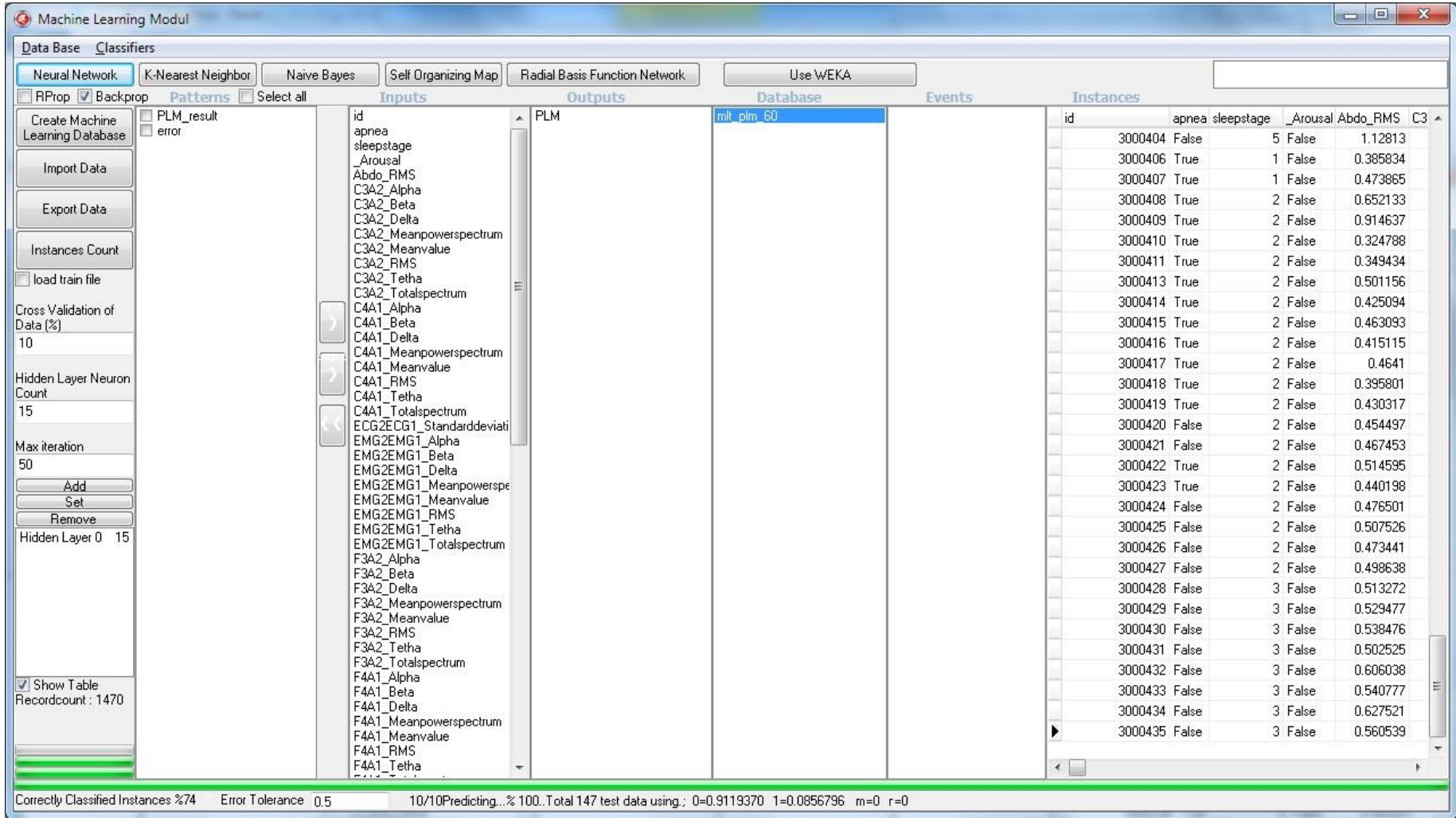
Her makine öğrenmesinin kendine özgü parametreleri solda bulunan (Şekil 3.13) paneldeki bileşenlerden veya çalışma esnasındaki çıkan seçeneklerden değiştirilebilir. Veriler çapraz doğrulama (cross-validation) yöntemi kullanılarak test edilir. Böylece tablodaki tüm veriler hem eğitim hem de test verisi olarak kullanılır. Çapraz doğrulamada kullanılacak olan veri yüzdesi ilgi text kutusundan değiştirilebilir.

Modül her tabloyu, öğrenme modelini ve sonuçlarını verilere özel bir isim oluşturarak saklayabilmektedir. Ayrıca bu sonuçları istatistiksel olarak karşılaştırabilmektedir. Saklanan bu öğrenme modellerinin, gerçek zamanlı bir sistemde kullanılması hedeflenmektedir. Örnek bir dosya görüntüsü Şekil 3.12’de gösterilmektedir. Şekildeki CTD dosyaları algoritmanın öğrenme modelinin sakladığı dosya biçimleridir. Dosyalar her bir veritabanı, tablo, makine öğrenmesi algoritması ve çapraz doğrulamanın bölütü için ayrı olarak isimlendirilerek saklanmaktadır.



Şekil 3.12 Makine Öğrenmesi Modülünün oluşturduğu örnek dosyalar

Şekil 3.13’de Makine Öğrenmesi Modülünün örnek bir arayüzü gösterilmektedir. Modülde tüm öznitelikler *pattern* etiketinde yer almaktadır. *Pattern* etiketindeki öznitelikler isteğe bağlı olarak ya da ‘*select all*’ seçenek kutusu ile seçilebilmektedir. Örnek şekilde bacak haricindeki tüm öznitelikler önce *input* listesine PLM ise *output* listesine taşınmıştır.



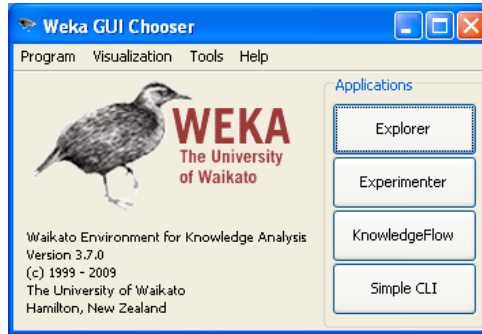
Şekil 3.13 Makine Öğrenmesi Modülünün ekran görüntüsü

Veriler modül dışında bir yazılımda kullanılmak istenildiğinde *Export Data* butonu ile farklı dosya biçimlerine aktarılır. Bu dışa aktarma işlemi sırasında aktarılacak dosya türüne göre farklı giriş kutularına parametreler girilir. Verilerin tamamının aktarımı yerine filtrelenerek de aktarım yapılabilir. 8 saat süren bir hasta kaydında epok uzunluğu 30 sn seçildiğinde 960 adet epok oluşmaktadır. Örneğin çıkış olarak seçilen öznitelik 0,1 (evet, hayır veya var-yok) şeklinde ve 960 epokta sadece 50 Evet, 910 Hayır vardır. Eğer evet veya hayırların hepsi aktarılmak istenmez ise istenilen değer kadar rastgele seçilerek diğer seçilmeyenler aktarımda hariç tutulabilir. Bu işlem tüm hasta kayıtlarında ya da her hastaya özel alt ve üst sınırlar tanımlanarak da yapılabilir.

Program ayrıca seçilen veritabanındaki tüm makine öğrenmesi tablolarını sırası ile WEKA yazılımına giriş verisi olarak vermektedir. Modül WEKA’da var olan algoritmalarından seçilmiş algoritma ve parametreleri ile beraber sıralı işlem olarak çalıştırılarak, istendiğinde sonuçlarını analize özel bir dosya olarak saklayıp karşılaştırma yapabilmektedir.

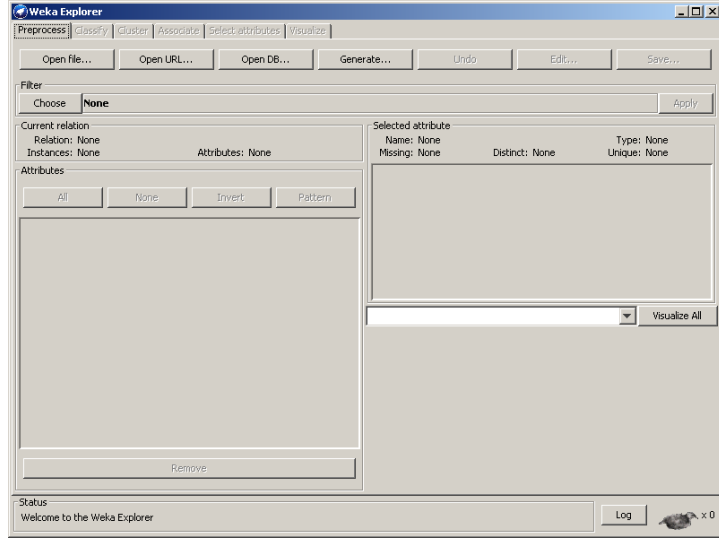
3.6.1 WEKA

WEKA makine öğrenme algoritmalarının ve veri ön işleme araçlarının bir araya getirildiği açık kaynak kodlu bir veri madenciliği programıdır. Yeni Zelanda’nın Waikato Üniversitesinde geliştirilmiş ücretsiz bir yazılımdır. Yazılım, Java yazılım dili ile geliştirilmiştir. Büyük veya dağıtık veri tabanlarında kullanılabilir. WEKA ile verinin hazırlanması, sınıflama, kümeleme, birliktelik analizi, nitelik değerlerinin seçilmesi yapılabilir. Arff, csv, c45 biçimindeki dosyalar kullanılabilir [115]. WEKA programının arayüzü Şekil 3.14’de görülmektedir.



Şekil 3.14 WEKA’da Applications Menüsü

WEKA çalıştırıldıktan sonra Şekil 3.14’de görüldüğü gibi, Application menüsü açılmaktadır. Bu menüde çalışma modları listelenmektedir. Simple CLI modu komut modunda çalışmaktadır. Explorer modu projeyi görsel ortamda adım adım gerçekleştirmeyi sağlar. KnowledgeFlow modu ise projeyi sürükle bırak yöntemiyle gerçekleştirmektedir [116]. Çalışmada kullandığımız explorer modunun arayüzü Şekil 3.14’de gösterilmektedir.



Şekil 3.15 WEKA’da veri seçimi

Şekil 3.15’deki arayüzde çalışılacak verilerin seçilmesi, bu veriler üzerinde temizleme ve dönüştürme gibi işlemler yapılır.

WEKA’da Arff, Csv, C4.5 formatında dosyalar işlenebilmektedir. WEKA’nın içerisinde Veri İşleme, Veri Sınıflandırma, Veri Kümeleme, Veri İlişkilendirme gibi özellikler vardır.

Yukarıda anlatılan adımlardan sonraki adımda çalışmanın amacına göre açılan Sınıflandırma, Kümeleme, İlişkilendirme gibi özellikler seçilerek uygun algoritmalar seçilerek veri kümeleri ile sınıflandırmalar yapılabilir [117].

4 PSG VERİLERİ ÜZERİNE UYGULAMA

Bölüm 3.1’de özniteliklerin elde edilmesi ayrıntılı olarak anlatılmıştı. Elde edilen öznitelikler Tablo 3.2’de gösterilmiştir. Bu öznitelikler makine öğrenmesinin farklı algoritmaları ile performans bakımından analiz edilerek karşılaştırılmıştır.

Veri sınıflama yöntemlerinin uygulanması aşamasında öncelikle WEKA, YALE olmak üzere bazı yazılım araçları incelenmiştir. YALE yazılımında izin verilen yöntem sayıları sınırlı sayıda olduğundan WEKA yazılım aracı tercih edilmiştir.

WEKA, sınıflandırıcı makine öğrenmesi tekniklerini içeren bir analiz programıdır. WEKA ‘arff’(Özellik-İlişki Dosya Biçimi) dosya formatındaki verileri analiz edebilmektedir. Çalışmada kullanılan veriler geliştirdiğimiz PSG analiz yazılımı ile ‘arff’ dosya formatına dönüştürülmektedir. WEKA programına yüklenen veriler 77 adet nitelik (Attributes) içermektedir. Aşağıda Şekil 4.1’de ‘arff’ dosya formatındaki bazı nitelikler ve nitelik sayıları gösterilmiştir.

```
=== Run information ===  
  
Scheme:weka.classifiers.functions.Logistic -R 1.0E-8 -M -1  
Relation:      mlt_plm_5  
Instances:     52602  
Attributes:    77  => Nitelik sayısı  
apnea  
sleepstage  
Abdo_RMS  
Airflow2_RMS  
C3A2_Alpha  
C3A2_Beta  
C3A2_Delta  
C3A2_Meanpowerspectrum  
C3A2_Meanvalue  
C3A2_RMS  
C3A2_Tetha  
C3A2_Totalspectrum  => Nitelikler
```

Şekil 4.1 ‘Arff’ dosya formatındaki niteliklerin bazıları ve niteliklerin sayısı

4.1 Modelin Kurulması

Veri madenciliği tekniklerini 5 ana grupta toplamak mümkündür. Bu gruplar sırasıyla; Sınıflama, Kümeleme, Regresyon Modelleri, Birliktelik Kuralı ve Ardışık Zamanlı Örüntülerdir. Tahmin etmede yararlanılan veri madenciliği yöntemleri içerisinde yaygın olarak kullanılan yöntemler, sınıflama ve regresyon modelleridir [118].

Tanımlanan problem için en uygun modelin bulunabilmesi, olabildiğince çok sayıda modelin kurularak denenmesi ile mümkündür. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılincaya kadar yinelenen bir süreçtir [119].

Tez çalışmasında sınıflandırma modellerine yer verilmiştir. Mevcut durumu değerlendirmek için sınıflandırma algoritmalarının tümü denenerek analiz edilmiştir. En yüksek performans gösteren yedi algoritmanın sonuçları karşılaştırılarak sıralanmıştır.

4.1.1 Sınıflandırma Başarımı Ölçüm Metrikleri

Genel olarak sınıflandırma problemi farklı akıllı sistem teknikleriyle ele alınır. Her sınıflandırma probleminin çözümü daha önce o problemi ele alan çözüm algoritmalarıyla veya aynı çalışma içindeki diğer algoritmaların başarımlarıyla belli bir standarda uygun olarak kıyaslanmalıdır. Her algoritma belli parametrelere bağlı olarak belirli sonuçlar ürettiğinden tüm sonuçları değerlendirmek için güvenilir karşılaştırma veya başarımlar ölçüm metriklerine ihtiyaç duyulmaktadır. Deneysel çalışma sonuçlarının istatistiksel sağlaması, sonuçlarla ilgili çıkarımların doğruluğu için önem taşır [120].

Öğreticili algoritmanın kullanıldığı çalışmalarda aynı problemin çözümünde kullanılan algoritmalar arasındaki sınıflandırma başarımlarının kıyaslanması için literatürde değişik yöntemler önerilse de en yaygın olan sınıflandırma başarımlar ölçüm metrikleri; genel doğruluk (Accuracy, ACC), f-ölçümü (F-measure, Fm), Matthew'un Korelasyon Katsayısı (Matthew's Correlation Coefficient, MCC), ROC eğrisi (Receiver Operating Characteristics Curve, ROC Curve), ROC eğrisi altında kalan alan (Area Under the ROC Curve, AUC), ortalama hata kareleri (Mean Squared Error, MSE), ortalama hata karelerinin kökü (Root Mean Squared Error, RMSE), Kappa Hatası

(Kappa's Error, KE) ve k-katlı apraz saęlama (K-fold Cross Validation) Őeklinindedir [121].

Model kurulurken verilerin belli bir kısmı eęitilmiŐ (training), eęitilen verilerin oluŐturduęu rntler kullanılarak geri kalan veriler test edilmiŐtir. WEKA sınıflandırma algoritmaları kullanarak, test veri sınıflarının tahmini iŐlemleri yapılmıŐtır. Makine ęrenmesi sınıflandırma algoritmalarının, performans karŐılaŐtırması Doęrululuk (Accuracy), RMSE (Root mean squared error), P [Hassasiyet(Precision)], ROC [AUC (Receiver Operating Characteristic)], TP (True Pozitif), FP (False Pozitif), TN (True Negatif), FN (False Negatif) deęerleri ile yapılmıŐtır.

Sınıflandırma algoritmalarının performanslarını deęerlendirmek amacıyla, znelik verilerimize uygun tm algoritmalar test edilmiŐtir. WEKA sınıflandırma algoritmalarını test etme aŐamasında, zneliklerin bazı algoritmalarla iŐlenmeye uygun olmadığı grlmŐtr. Bu algoritmalar performans deęerlendirme iŐlemine dahil edilmemiŐtir.

alıŐmada kullanılan tm sınıflandırıcılar iin test seeneęi olarak 10 kat aprazlama (10 fold cross-validation) teknięi kullanılmıŐtır. Bu yntemde veri kmesi 10 eŐit paraya blnr. 10 eŐit paradan 1 tanesi test iin, geriye kalan 9 tanesi eęitim iin kullanılır [15]. Test verileri 10 kat aprazlanmış veri kmesi zerine WEKA'da bulunan sınıflandırma algoritmalarıyla test edilmiŐtir.

4.2 Modelin Deęerlendirilmesi

Bir model deęerlendirilirken gz nnde bulundurulan baŐarı ltleri sınıflandırma modeli altındaki "rules.JRip" algoritmasının sonuları aŐaęıdaki Őekilde gsterilmiŐtir. Bu algoritmanın sonuları rnek alınarak performans deęerlendirme kriterleri anlatılmıŐtır.

```

=== Run information ===

Scheme:weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1
Relation:      mlt_plm_60
Instances:     9286
Attributes:    77

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      6460           69.5671
%
Incorrectly Classified Instances    2826           30.4329
%
Kappa statistic                    0.3913
K&B Relative Info Score            219674.3573 %
K&B Information Score              2196.7436 bits    0.2366
bits/instance
Class complexity | order 0         9286.0011 bits    1
bits/instance
Class complexity | scheme          9371.5127 bits    1.0092
bits/instance
Complexity improvement (Sf)        -85.5116 bits    -0.0092
bits/instance
Mean absolute error                0.3967
Root mean squared error            0.4611
Relative absolute error            79.3318 %
Root relative squared error        92.2247 %
Total Number of Instances          9286

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure
ROC Area  Class
0.717     0         0.657     0.266     0.712     0.657     0.683
0.717     1         0.734     0.343     0.682     0.734     0.707
Weighted Avg. 0.696     0.304     0.697     0.696     0.695
0.717

=== Confusion Matrix ===

      a    b  <-- classified as
3051 1592 |    a = 0
1234 3409 |    b = 1

```

Şekil 4.2 ‘rules.JRip’ algoritmasının sonuçları

‘Rules.JRip’ algoritmasının sonuçlarına göre TN (True Negatif) = 3051, FN (False Negatif) = 1592, FP (False Pozitif) = 1234, TP (True Pozitif) = 3409 değerleri ‘Confusion Matrix’ sonuçlarından elde edilmiştir. Toplam, doğruluk, hata oranı, hassasiyet, TNR, TPN değerlerinin denklemleri aşağıda gösterilmiştir. Denklemlerdeki tüm değerler ‘Confusion Matrix’ kullanılarak hesaplanmaktadır [122].

4.2.1 Karışıklık Matrisi

Confusion Matrix (CM) bir akıllı sistem algoritmasına ait sınıflandırma başarımı hakkında bütüncül bir yaklaşımla bilgi veren bir matris modelidir. CM yapısı sınıflandırıcının başarımı ve yapılan test ile yakından ilgilidir. Bu şekilde bir

sınıflandırma başarımının CM olarak türetilmesi, deneysel çalışma ile ilgili tüm başarım metriklerinin kolayca hesaplanmasını mümkün kılar [123]. Bir CM yapısal olarak aşağıdaki denklemde olduğu gibi ifade edilir. CM’de her elemanın sınıflandırma başarımı ile ilgili özel bir anlamı vardır. Bu elemanlar diğer başarım metriklerinin tanımlanmasında referans olarak kullanılırlar.

$$CM = \begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \quad (4-1)$$

Yukarıdaki matriste gösterildiği gibi dört elemana sahip olan CM’de her eleman, gerçek pozitif (True Positive, TP), yalancı pozitif (False Positive, FN), yalancı negatif (False Negative, FN) ve gerçek negatif (True Negative, TN) şeklinde tanımlanır. Daha açık bir ifadeyle matrisin her bir elemanı bir deneysel gözlemin sonucuna göre gerçek değer ile tahmin edilen değer arasındaki ilişkiyi gösterir [124].

4.2.2 İşlem Karakteristiği Eğrisi (ROC-AUC)

Hastalığın doğru ve erken teşhisi, hastalığın hasta üzerindeki etkisini önemli ölçüde değiştirmektedir. Bu sebeple hastalık teşhis yöntemleri tıp alanında ciddi bir araştırma konusu olmuştur. Yapılan çalışmalarda genellikle, teşhis yöntemlerinin performansı, test kümesindeki doğru sınıflandırma oranı ile ölçülmekte ve elde edilen bu oran doğru teşhis oranı olarak verilmektedir. Sonucunda hayati tehlikeler oluşabilecek türdeki hastalıkların yanlış teşhisinde, kaç hasta kişiye hasta, kaç sağlam kişiye sağlam teşhisi konulduğunun yanı sıra kaç hasta kişiye sağlam teşhisi ve kaç sağlam kişiye hasta teşhisi konulduğunun da çok büyük önemi vardır [125].

ROC adı, bir radar terimi olan ve doğru radyo sinyallerini gürültüden ayırt etmeye yarayan Receiver Operating Characteristic’ten gelmektedir.

Bu yöntem,

1. Testin ayırt etme gücünün belirlenmesine,
2. Çeşitli testlerin etkinliklerinin kıyaslanmasına,
3. Uygun pozitiflik eşiğinin belirlenmesine,
4. Laboratuvar sonuçlarının kalitesinin izlenmesine,
5. Uygulayıcının gelişiminin izlenmesine ve
6. Farklı uygulayıcıların etkinliklerinin kıyaslanmasına olanak sağlar [125].

Teşhislerin doğruluğunun belirlenebilmesi için duyarlılık (sensitivity) ve belirlilik (specificity) kavramlarından yararlanır. Bu kavramların tanımları; TP (true positive), TN (true negative), FP (false positive), FN (false negative) sırasıyla hasta kişiye hasta teşhisi konması, sağlam olan kişiye sağlam teşhisinin konması, sağlam olan kişiye hasta teşhisi konması ve hasta olan kişiye sağlam teşhisi konması olmak aşağıdaki üzere ifadeleriyle verilir [126].

$$\text{Total(Toplam)} = (TP + TN + FP + FN) \quad (4-2)$$

$$\text{Total(Toplam)} = (3409 + 3051 + 1234 + 1592) = 9286$$

$$\text{Doğruluk(Accuracy)} = (TP + TN) / (\text{Total}) \quad (4-3)$$

$$\text{Doğruluk(Accuracy)} = (TP + TN) / \text{Total} = (3409 + 3051) / 9286 = \%69.5671$$

$$\text{Hataoranı} = (FP + FN) / \text{Total} \quad (4-4)$$

$$\text{Hataoranı} = (FP + FN) / (\text{Total}) = (1234 + 1592) / 9286 = 2826 / 9286 = \%30.4329$$

$$0 \text{ için Hassasiyet(Precision)} = (TN) / (TN + FN) \quad (4-5)$$

$$1 \text{ için Hassasiyet(Precision)} = TP / (TP + FP) \quad (4-6)$$

$$0 \text{ için Hassasiyet(Precision)} = TN / (TN + FN) = 3051 / 4643 = 0.657$$

$$1 \text{ için Hassasiyet(Precision)} = TP / (TP + FP) = 3409 / 4643 = 0.734$$

$$0 \text{ için Recall TNR(True Negative Rate)} = \text{TNR} = TN / (TN + FP) \quad (4-7)$$

$$1 \text{ için Recall TPR(True Positive Rate)} = \text{TPR} = TP / (TP + FN) \quad (4-8)$$

$$\text{TNR} = TN / (TN + FP) = 3051 / 4285 = 0.712$$

$$\text{TPR} = TP / (TP + FN) = 3409 / 5001 = 0.682$$

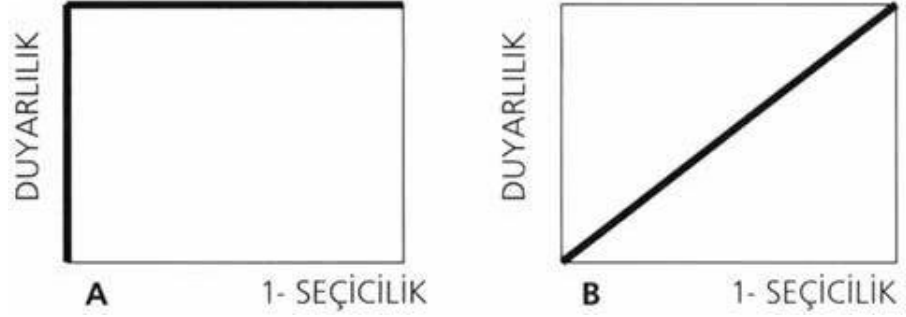
$$0 \text{ için F1 - Score} = (2 \times \text{Precision} \times \text{recall}) / (\text{Precision} + \text{recall}) \quad (4-9)$$

$$1 \text{ için F1 - Score} = (2 \times \text{Precision} \times \text{recall}) / (\text{Precision} + \text{recall}) \quad (4-10)$$

$$\text{F1-Score for 0} = (2 \times \text{Precision} \times \text{recall}) / (\text{Precision} + \text{recall}) = 0.935568 / 1.369 = 0.683$$

$$\text{F1-Score for 1} = (2 \times \text{Precision} \times \text{recall}) / (\text{Precision} + \text{recall}) = 1.001176 / 1.416 = 0.707$$

İşlem karakteristiği eğrisi, koordinat sisteminin ordinatına duyarlılık ve absisine ise yanlış pozitiflik (1-Seçicilik) konularak çizdirilir. Tanı testi ne kadar iyi ise eğri o kadar yukarıya ve sola doğru kayar (Şekil 4.3).



Şekil 4.3 İdeal (A) ve kötü performans (B) gösteren iki testin ROC eğrisi

Böyle bir eğri çizildikten sonra, eğrinin altında kalan alan hesaplanabilir (Area Under Curve–AUC). AUC değeri ne kadar yüksek ise, tanı testi o kadar iyi bir ayırım yeteneğine sahiptir. Örneğin 0.975 üzerindeki AUC değerleri, mükemmel sayılmaktadır. AUC'nin 0.5'den farklı olup olmadığı istatistiksel olarak test edilmeli ve buna ait p değeri de birlikte verilmelidir. İki farklı teste ait AUC değerleri arasındaki fark istatistiksel olarak anlamlı ise AUC'si büyük olan test daha fazla ayırım gücüne sahip demektir [127].

Makine öğrenmesi sınıflandırma algoritmalarının, performans karşılaştırması Doğruluk (Accuracy), RMSE (Root mean squared error), P (Hassasiyet(Precision)), ROC [AUC (Receiver Operating Characteristic)], TP (True Pozitif), FP (False Pozitif), TN (True Negatif), FN False Negatif) değerleri ile analiz edilmiştir. Bu değerlere göre makine öğrenmesi algoritmaları 5 saniyelik ve 60 saniyelik epoklar için hesaplanmış ve Tablo 4.1, Tablo 4.2'de sıralanmıştır.

Tablo 4.1 60 sn'lik epok süreleri ile makine öğrenmesi algoritmalarının performans analizi

Algoritma Adı	PLM Epok 60 sn.							
	D %	RMS	P	ROC (AUC)	TP	FP	TN	FN
classifiers.lazy.IBk	79.00	0.4652		0.785	3686	957	3590	1053
trees.J48	71.86	0.515		0.721	3369	1274	3304	1339
functions.MultilayerPerceptron	70.80	0.4806		0.764	3242	1401	3332	1311
rules.JRip	69.57	0.4611		0.717	3409	1234	3051	1592
rules.DTNB	66.97	0.4615		0.733	3148	1495	3070	1573
rules.DecisionTable	64.60	0.4721		0.691	2988	1655	3010	1633
functions.Logistic	63.34	0.4741		0.685	2737	1906	3144	1499

60 sn'lik epok uzunlukları ile elde edilen makine öğrenmesi algoritmalarının sonuçları yukarıda Tablo 4.1'de gösterilmiştir. Bu sonuçlara göre en yüksek doğruluk yüzdesini % 79 ile '**classifiers.lazy.IBk**' algoritması, en düşük doğruluk yüzdesini ise % 63.34 ile '**functions.Logistic**' algoritması vermiştir.

5 sn'lik epok uzunlukları ile elde edilen makine öğrenmesi algoritmalarının sonuçları aşağıda Tablo 4.2'de gösterilmiştir.

Tablo 4.2 5 sn'lik epok süreleri ile makine öğrenmesi algoritmalarının performans analizi

Algoritma	PLM Epok 5 sn.							
	D %	RMS	P	ROC (AUC)	TP	FP	TN	FN
rules.JRip	63.93	0.4792		0.657	16359	9942	17268	9033
classifiers.lazy.IBk	63.82	0.6015		0.638	16834	9467	16733	9568
functions.MultilayerPerceptron	63.15	0.4822		0.68	16909	9392	16306	9995
functions.Logistic	62.87	0.4762		0.676	14929	11372	18142	8159
rules.DTNB	62.53	0.4819		0.672	15863	10438	17024	9277
rules.DecisionTable	62.34	0.4776		0.671	15270	11031	17519	8782
trees.J48	62.00	0.5977		0.619	16349	9952	16266	1005

Tablo 4.2'de sonuçlara göre en yüksek doğruluk yüzdesini % 63.93 ile '**rules.JRip**' algoritması, en düşük doğruluk yüzdesini ise % 62 ile '**trees.J48**' algoritması vermiştir.

4.3 Uygulama Sonuçları

UPHB, bacak EMG'si öznitelikleri haricindeki diğer tüm öznitelikler kullanılarak makine öğrenmesi algoritmalarıyla otomatik olarak skorlanmıştır. Makine öğrenmesi sınıflandırma algoritmalarının, performans karşılaştırması Doğruluk (Accuracy), RMSE (Root mean squared error), P [Hassasiyet(Precision)], ROC (Receiver Operating Characteristic), TP (True Pozitif), FP (False Pozitif), TN (True Negatif), FN (False Negatif) değerleri ile analiz edilmiştir.

Tablo 4.1 ve Tablo 4.2'de bacak EMG'si kanalları hariç diğer kanallar kullanılarak çeşitli makine öğrenmesi algoritmalarıyla sınıflandırma başarı yüzdeleri, sıralı olarak gösterilmiştir.

60 sn'lik epoklar kullanılarak sınıflandırma başarı yüzdesi baz alınarak elde edilen sonuçlara göre 'classifiers.lazy.IBk' algoritması %79.00 doğruluk yüzdesi ile en başarılı sınıflandırma yüzdesine sahiptir.

Benzer şekilde 5 sn'lik epoklar kullanılarak 'rules.JRip' algoritması % 63.93 başarı yüzdesiyle en başarılı doğruluk yüzdesi ile sınıflandırma yapmıştır. 5 sn'lik epok süresi kullanılarak elde edilen sonuçlara göre, 'classifiers.lazy.IBk' algoritması 2. yüksek sınıflandırma yüzdesine sahiptir. Başarılı sınıflandırma yüzdelerine göre 5, 60 sn'lik epok sürelerine göre 'classifiers.lazy.IBk' algoritması 1. ve 2. sırada yer almasından dolayı en başarılı algoritma olduğu sonucuna varılmıştır.

PSG analiz yazılımı, PSG kaydındaki sinyallerin özniteliklerini 2, 5, 30, 60 sn'lik epok (bölüt) uzunlukları ile elde edebilmesine olanak sağlamaktadır. Tez sonuçları elde etme aşamasında, yapmış olduğumuz ön çalışmalar ve deneyler esnasında 2, 5, 30, 60 sn'lik epok uzunlukları kullanılarak elde edilen öznitelikleri sınıflandırma işlemine tabi tuttuğumuzda, seçilen epok süreleri arttıkça sınıflandırma yüzdelerinin arttığı sonucuna varılmıştır. Sınıflandırma yüzdeleri bakımından 60 saniyelik epok uzunluğuna sahip öznitelikler kullanılarak elde edilen sonuçların, 5 saniyelik epok uzunluğuna sahip öznitelikler kullanılarak elde edilen sonuçlardan daha iyi sonuçlar verdiği gözlemlenmiştir.

Aynı şekilde 5 saniyelik epok uzunluğuna sahip öznitelikler kullanılarak elde edilen sonuçların, 2 saniyelik epok uzunluğuna sahip öznitelikler kullanılarak elde edilen sonuçlardan daha iyi sonuçlar verdiği gözlemlenmiştir. Bu nedenle tez

sonuçlarının değerlendirilmesinde, 60 ve 5 saniyelik epok uzunlukları kullanılarak elde edilen sınıflandırma algoritmalarının sonuçlarına yer verilmiştir.

4.4 Uygulama Sonuçlarının Değerlendirilmesi

Çalışmanın zayıf yönleri olmasıyla birlikte, güçlü yönleri fazladır. Sonuçların elde edilmesinde tanımlayıcı (basit istatistik yöntemleri) istatistik yöntemleri dışında, ileri istatistik yöntemleri kullanılmıştır. PSG verilerinin analizinde üç farklı yöntem kullanılmıştır. Bu yöntemler WEKA programı, İstatiksel yöntem ve çalışmada geliştirdiğimiz PSG Analiz Programında kullanılan yöntemlerdir.

Çalışmada 153 gerçek hasta verisi kullanılmıştır. Sadece bir hasta için 2 saniyelik epok uzunlukları seçildiğinde 14400 epok sayısı elde edilmektedir. Tüm hastalar için 2203200 epok sayısına ulaşılmaktadır. Yüksek miktarda veri kullanmak, yapılan analizlerden daha anlamlı sonuçlar elde edilmesini sağlamıştır.

PSG analiz yazılımı istenilen epok süresi seçimine olanak sağlamaktadır. Böylece istenilen uzunluktaki epok sürelerine sahip, öznelikler elde edilebilmektedir.

DSP ve Makine Öğrenmesi Modülünde, verilerin (özneliklerin) elde edilmesinde kullanılan araç bileşenleri, ücretli ve lisanslı yazılımlardır.

Yazılım modüler yapısından dolayı, farklı sinyal analizlerinde de kullanılabilir (Epilepsi, EEG sinyallerinin analizinde).

Çalışmanın zayıf yönleri ise, otomatik skorlama işlemi sırasında çok az karşılaşılsa da hatalar olmaktadır. Bu hatalar nadiren, uzman hekimlerin gözünden kaçabilmektedir. Yanlış bir skorlama verisi analizi olumsuz yönde etkileyebilmektedir.

PSG kayıtlarının bazı epoklarında artefaktlar bulunmaktadır. Bu artefaktlar hastanın vücudundan (göz hareketleri, EKG) ve dış etkenlerden (Enstrümental, Elektrot) kaynaklanmaktadır. Artefaktlar PSG kaydındaki kanallarda, çeşitli gürültülere neden olmaktadır. Kanallar üzerinde oluşan artefaktlar, PSG Analiz Programının Artefakt Modülü ile temizlenmiştir. Artefaktlar yoğun olduğunda, kanallardaki sinyallerin doğal yapısı bozulmaktadır. Bu durum analiz sonucunu olumsuz yönde etkilemektedir.

Artefakt konusunda daha ayrıntılı bir tespit ve temizleme yöntemi geliştirilebilirse, artefaktlardan dolayı oluşan olumsuz etkiler en aza indirilebilir.

Yapılan tez çalışmasında anlamlı sonuçlar elde edilmiştir. Makine öğrenmesi algoritmalarından 'lazy.Ibk' sınıflandırma algoritması ile bacak EMG sinyali kullanılmadan UPHB % 79 oranında sınıflandırılmıştır. Makine öğrenmesinde kullanılan algoritmalar, aynı veri setleri kullanılarak birden çok kez tekrar edilmiştir. Bu tekrarlarla elde edilen sonuçların birbirine çok yakın olduğu gözlemlenmiştir. Bu gözlemler sonuçların rastlantısal olmadığını göstermektedir.

4.5 Öneriler

PSG kaydı insan vücudunun başından ayaklarına kadar çok sayıda elektrot veya sensör yerleştirilmesi ile gerçekleştirilmektedir. Eğer UPHB, sağ ve sol bacağa yerleştirilen elektrotlar haricinde çekim için kullanılan diğer sensör veya elektrotlardan alınan sinyaller aracılığı ile yapılabilirse bacaklara bağlanan elektrotlardan tasarruf edilmiş olacaktır.

Kayıt esnasında sinyallerin alınacağı elektrotlar veya sensörler, hastanın vücuduna, PSG kaydı öncesinde uyku teknisyeni tarafından yerleştirilir. Bu işlem yaklaşık olarak 30 dakika sürmektedir. Elektrotları yerleştirmek oldukça zahmetli bir iştir. Elektrotların ve sensörlerin yerleştirilmesi esnasında hasta beklemektedir. Eğer hastanın vücuduna yerleştirilecek elektrot veya sensör sayısı azaltılırsa, uyku teknisyeninin iş yükü de azalmış olacaktır. Ayrıca elektrot ve sensör yerleştirme süresi kısalmış olacağından hastanın bekleme süresi de azalacaktır. Hasta gece uyandığında vücudundaki elektrot ve sensör sayısı azaldığında daha rahat hareket edebilecektir.

Sağ ve sol bacağa bağlanan elektrotlar saniyede 128 örnekleme sayısı ile bilgisayara kayıt edilmektedir. Eğer bu iki elektrot kullanılmayacak olursa depolama birimlerine kayıt edilen veri sayısı azalacağı için depolama birimi daha tasarruflu kullanılmış olacaktır.

Hekimler uyku hastalıklarını, 6-8 saatlik PSG kaydının her bir kanalını göz ile tek tek incelemek suretiyle teşhis etmektedirler. Bu işleme skorlama denmektedir.

Skorlama işlemi zaman isteyen bir iştir. Eğer PSG kaydındaki kanal sayısında bir azalma olursa, hekimin gözle inceleyeceği kanal sayısı azalacağı için skorlama işlemi daha az sürede gerçekleşecektir.

PSG kaydında iki bacak EMG'si için, dört adet kablo ve 2 adet elektrot kullanılmaktadır. Eğer bacak EMG'si kanalları kullanılmazsa kablo ve elektrot tasarrufu sağlanacaktır.

KAYNAKLAR

- [1] http://www.ninds.nih.gov/disorders/brain_basics/understanding_sleep.htm, Erişim Tarihi: 23 Temmuz 2013.
- [2] Karadeniz D, Ondze B, Besset A, Billiard M., Are Periodic Leg Movements during Sleep (UPHB) Responsible for Sleep Distruption in İnsomnia Patients?, Eur. J Neurol, 54:502, 2000.
- [3] Thorpy M.J., Rochester Mina, The International Classification of Sleep Disorders, Diagnostic and Coding Manual and the Diagnostic Classification Steering Committee. American Sleep Disorders Association, 1990.
- [4] American Sleep Disorder Association, Recording and Scoring Leg Movements, Atlas Task Force of the American Sleep Disorders Association, Sleep 1993; 16:748–59.
- [5] Magdolna Hornyak, Bernd Feige, Dieter Riemann, Ulrich Voderholzer, Periodic Leg Movements in Sleep and Periodic Limb Movement Disorder, Prevalence, Clinical and Treatment. Sleep Medicine Reviews (2006) 10, 169–177.
- [6] Demir, Ömer, EEG Dalgalarının Wavelet (Dalgacık) Dönüşümü ile Değerlendirilmesi, Yüksek Lisans Tezi, Elektrik-Elektronik Mühendisliği Anabilim Dalı, 2008.
- [7] Mutlu, Sunay, EEG Sinyallerine Yapay Zeka Tekniklerinin Uygulanması, Yüksek Lisans Tezi, Gazi Üniversitesi, Bilişim Enstitüsü, 2008.
- [8] Ç., Elmas, Yapay Sinir Ağları, Ankara: Seçkin Yayıncılık, 2003.
- [9] Miner, N.E, An Introduction to Wavelet Theory and Analysis, Sandia Report, NM.
- [10] Batar, Hatice, EEG İşaretlerinin Dalgacık Analiz Yöntemleri Kullanılarak Yapay Sinir Ağları ile Sınıflandırılması, Kahramanmaraş: Yüksek Lisans Tezi,KSÜ Fen Bilimleri Enstitüsü 102s, 2005.
- [11] Altınbaş, Atılay, EMG Sinyallerinin Kısa Zamanlı Fourier Dönüşümü ve Dalgacık Dönüşümü Kullanarak Analiz Edilmesi, Yüksek Lisans Tezi, 2007.
- [12] Chan K., Lee T., Sample P.A., Goldbaum M.H., Weinreb R.N., Sejnowski T.J., Comparison of Machine Learning and Traditional Classifiers in Glaucoma Diagnosis, cilt 49, IEEE Transactions on Biomedical Engineering, Part. 49, No. 9, 963-973, pp. 963-973.
- [13] S. Daniş, Prostat Kanseri Olma Olasılığı Olan Biyopsi Öncesi Hastalarda Yanlış Tanıyı Azaltmak için Veri Madenciliği Uygulaması, Yüksek Lisans Tezi, 2003.
- [14] M. Ş. Toprak, Tıp Alanında Makine Öğrenmesi Uygulaması, Yüksek Lisans Tezi, 2004.
- [15] Y. Uzun, Tıbbi Veriler Üzerinde Makine Öğrenme Algoritmaları ve Bulanık Mantık ile Kurallar Öğrenme, Yüksek Lisans Tezi, 2005.
- [16] Eliash Z. B., Wollstein G., Chu J., Ramsey J.D., Glymour C., Noecker R.J.,

- Ishikawa H., Schuman J.S., Optical Coherence Tomography Machine Learning Classifiers for Glaucoma Detection: A Preliminary Study, cilt 46, Investigative Ophthalmology & Visual Science, Pittsburgh, Part. 46, No. 11, 4147-4152, pp. 4147-4152.
- [17] İ. Z. Gökbay, Meme Kanserinde Otomatik Öğrenme Teknikleri, Yüksek Lisans Tezi, 2007.
- [18] Bock R., Meier J., Michelson G., Ny'u L. G., Hornegger J., Classifying Glaucoma with Image-Based Features from Fundus Photographs, University of Erlangen-Nuremberg, Erlangen, 355–364, 2007.
- [19] M. Amasyalı, Yeni Makine Öğrenmesi Metotları ve İlaç Tasarımına Uygulamaları, Doktora Tezi, 2008.
- [20] Kahraman, Web-Tabanlı Uyarlanır Zeki Öğretim Sistemi Tasarımı ve Uygulaması, Doktora Tezi, 2009.
- [21] Çalış, Çağdaş, Veri Madenciliği Yöntemleri Kullanılarak Yüksek Tansiyon Hastalığı için İlaç Dozu Planlanması, Yüksek Lisans Tezi, 2010.
- [22] Özky, Aslı Uyar, Tüp Bebek Tedavi Sürecinde Yapay Öğrenme Yöntemleri: İmplantasyon ve Blastosist Gelişiminin Kestirimci Modellenmesi, Doktora Tezi, 2011.
- [23] M. Karabulut, Biyolojik Dizilimler Üzerinde Veri Madenciliği Teknikleri Kullanarak Transkripsiyon Faktörü Bağlanma Sitelerinin Tespiti, Doktora Tezi, 2011.
- [24] Ş. Hacıfendioğlu, Makine Öğrenmesi Yöntemleri ile Glokom Hastalığının Teşhisi, Konya: Yüksek Lisans Tezi, 2012.
- [25] Ü. C. Kumdereli, Tıp Bilişimi ve Veri Madenciliği Uygulamaları: EEG Sinyallerindeki Epileptiform Aktiviteye Veri Madenciliği Yöntemlerinin Uygulanması, Edirne: Yüksek Lisans Tezi, 2012.
- [26] A. Vahaplar, Bilişsel EEG Kayıtları Üzerinde Veri Madenciliği Uygulaması,, Doktora Tezi, 2012.
- [27] F. Yakar, Uyku Apne Sendromu Tanısında Gündüz Polisomnografinin Değeri, İstanbul: Uzmanlık Tezi, 2009.
- [28] Rechtschaffen A, Kales A (eds)., A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages In Human Objects, Los Angeles: Brain Research Institute, 1-13, 1973.
- [29] Conrad Iber, Sonia Ancoli, Andrew L.Chesson Jr., Stuart F.Quan, The AASM Manual for the Scoring of Sleep and Associated Events, Rules, Terminology And Technical Specifications. Sleep 2007, 2007.
- [30] Michael H. Silber, Lois E. Krahn, Timothy I. Morgenthaler, Sleep Medicine in Clinical Practice, Taylor&Francis, 2004.
- [31] Alpaslan Ersöz, Seral Özşen, Uyku EEG Sinyalinin Yapay Sinir Ağ Modeli İle Sınıflandırılması, Elektrik-Elektronik ve Bilgisayar Sempozyumu, 2011.
- [32] AARC-APT, Polysomnography Respir Care, AARC-APT (American Association of Respiratory Care-Association of Polysomnography Technologists) Clinical Practice Guideline, 1995.
- [33] Medicine, American Academy of Sleep, The AASM Manuel For The Scoring Of

- Sleep And Associated Events, American Academy of Sleep Medicine, 2007.
- [34] Okur HK., Polisomnografi Tanım ve Endikasyonları, Uyküder 2.Polisomnografi Sertifika Program Kurs Kitabı, 2008.
- [35] ED, Duygu, Sensörler ve Hastanın Kayıt için Hazırlanması.
- [36] Türkoğlu, Merve Yasemen, EEG Sinyallerinin Analizinde Performansı Yüksek Olan Dalgacık Tipinin Belirlenmesi, 2010: Yüksek Lisans Tezi, Elektrik-Elektronik Mühendisliği TOBB Ekonomi ve Teknoloji Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- [37] Ochoa, J.B., EEG Signal Classification for Brain. Computer Interface Applications., Ecole Polytechnique Federale De Luussanne.
- [38] Sanei S., Chambers, J.A., EEG Signal Processing, 2007.
- [39] Yazgan, E. ve Korkürek, M., Tıp Elektronikü, İstanbul Teknik Ünivertsitesi Yayınları, 386, 1996.
- [40] Sanei S., CHambers, J.A., EEG Signal Processing, 2007.
- [41] L., Karasulu, PSG'de Genel Prensipler, Kayıt Yöntemleri ve Kalibrasyon, Uyküder 2.Polisomnografi Sertifika Programı, Kurs Kitabı, 2008.
- [42] Korürek, Mehmet, Tıp Elektronikünde Tasarım İlkeleri, İ.T.Ü. Ofset Baskı Atölyesi, 975-561-081-2, İstanbul.
- [43] Rputchi, David and Norris, Micheal, Design and Development of Medical Electronic Instrumentatio, New Jersey: John Wiley & Sons, 0-471-676233, 2005.
- [44] Yüksel, Beyazıt Bestami, EKG İşaretlerinin Gömülü Sistem ile İzlenmesi, İstanbul: Yüksek Lisans Tezi, Elektronik Bilgisayar Eğitimi Anabilim Dalı Bilgisayar ve Kontrol Eğitimi Programı, 2011.
- [45] Levy P, Peppin JL, Wuyam B, Veale D., Respiratory Monitoring in Sleep Apnea Syndrome, Sleep 1992;15:S5-S8, 1992.
- [46] Yılmaz Hikmet, Çekimi Başlatma ve Monitorizasyon, 11. Uyku Tıbbı K ongresi, 6-10 Kasım 2010.
- [47] Hosselet JJ, Norman RG, Ayappa I, Detection of Flow Limitation Wit A Cannula/Pressure Transducer System, Sleep 2000;23:763-71, 2000.
- [48] Rechtschaffen A, Bergmann BM, Everson CA, et al., Sleep Deprivation in The Rat: X. Integration and Discussion f the findings, Sleep 1989;12:68-87, 1989.
- [49] Allen RP, Picchietti D, Hening WA, et al: , Restless legs syndrome:Diagnostic criteria, special considerations, and epidemiology., A report from the restless legs syndrome diagnosis and epidemiology workshop at the National Institutes of Health. Sleep Med 4:101-119, 2003.
- [50] Iber C, Ancoli-Israel S, Chesson AL, Quan S., The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification, Westchester, IL: American Academy of Sleep Medicine, 2007.
- [51] Çokçetin, Bahadır, PHP, MYSQL Tabanlı Uzaktan Eğitim Modülü Tasarımı, Yüksek Lisans Tezi, Elektrik-Elektronik Mühendisliği Anabilim Dalı, 2006.
- [52] Mintzer F., Filters for Distortion-Free Two Band Multirate Filter Banks, IEEE Trans. Acoust., Speech, and Signal Proc., 33(3), 626-630, 1985.
- [53] <http://users.rowan.edu/~polikar/WAVELETS/WTpart4.htm>, Erişim Tarihi: 29 Temmuz 2013.

- [54] Namba, M., Ishida, Y., Wavelet Transfrom Domain Blind Deconvolution, Signal Processing, 68, 119-124 p, 1998.
- [55] Koçyiğit Y., Çok Fonksiyonlu Kol Protezleri için Elektromiyogram İşleme Sistemi, İstanbul: Doktora Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2004.
- [56] Akay M., Wavelet Applications in Medicine, IEEE Spectrum, 34(5), 50-56, 1997.
- [57] Mallat S.G., Theory for Multiresolution Signal Decomposition: The Wavelet Representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), 674-693, 1989.
- [58] Ertürk S., Sayısal İşaret İşleme, İstanbul: Kocaeli Üniversitesi, Birsen Yayınevi, 2004.
- [59] Bergland, G. D., A Guide Tour of the Fast Fourier Transform, IEEE Spectrum, Vol. 6, pp. 41-55, 1969.
- [60] Misiti M., Misiti Y., Oppenheim G., Poggi J. M., Wavelet Toolbox for Use with MatLab, The MathWorks, 1996.
- [61] Haselsteiner, E. Pfurtscheller, G., Using Time-Dependent Neural Networks for EEG Classification, IEEE Trans Rehab Eng, 8: 457-63, 2000.
- [62] Roberts, R.A., Mullis, C.T., Digital Signal Processing, New York: Addison-Wesley Publishing Company, 85-102, 1987.
- [63] Erdoğan P, Pekçakar A, Dalgacık Dönüşümü ile EKG Sinyallerinin Özellik Çıkarımı ve Yapay Sinir Ağları ile Sınıflandırılması, 5. Uluslararası İleri Teknolojiler Sempozyumu (IATS'09), Karabük, Türkiye, 2009.
- [64] Küçük M., Ağralıoğlu N., Dalgacık Dönüşüm Tekniği Kullanılarak Hidrolojik Akım Serilerinin Modellenmesi, İTÜ Mühendislik, Dergisi 2006, S 69-80.
- [65] Ali Ustaoglu C., Ertunç H.M., Ocak H., Arıza Gözlemlemeye Yönelik Rulman Hata Tespit Algoritmalarının Karşılaştırmalı Analizi, Otomatik Kontrol Ulusal Toplantısı, 2008.
- [66] Serdaroglu A., Ertüzün A., Erçil A., Tekstil Kumaş İngelerinde Dalgacık Dönüşümleri ve Bağımsız Bileşen Analizi ile Hata Denetimi,, Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, 2006.
- [67] Ayaz E., Şeker S., Türkcan E., Barutçu B., Elektrik Motorlarında Akım İşaretlerini Kullanarak Dalgacık ve Spektral Analiz Yöntemlerinin Birleşimi ile Arıza Tanısı, Elektrik-Elektronik-Bilgisayar Mühendisliği 10. Ulusal Kongresi, 118-121, 2003.
- [68] Soltani, S., On The Use of The Wavelet Decomposition for Time Series Prediction, Neurocomputing, 2002.
- [69] I., Daubechies, The wavelet Transform, Time-Frequency Localization and Signal Analysis, IEEE Transactions on Information Theory, 36(5), 961-1005, 1990.
- [70] Karakoç, Ali Oktay, Kalman Filtresi ve Ayrık Dalgacık Dönüşümü Tekniği Kullanılarak EKG İşaretlerinin Üzerindeki Gürültülerin Temizlenmesi, İTÜ Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Elektronik ve Haberleşme Mühendisliği, 2010.
- [71] J. F. Scholl, J. R. Agre, L. P. Clare, and M. C. Gill, A Low Power Impulse Signal Classifier Using the Haar Wavelet Transform, in Proc. SPIE, Sensors, pp 136-

- 145, 1999.
- [72] Cuiwei Li, Chongxun Zheng, and Changfeng Tai , Detection of ECG Characteristic Points Using Wavelet Transforms IEEE Transactions On Biomedical Engineering, Vol. 42, No.14, January, pp. 21 – 28, 1995.
- [73] Yardımcı, Tuğrul, Makine Öğrenmesi Teknikleri İle Rss Besleme Yönetimi, Ankara: Yüksek Lisans Tezi, Bilgisayar Eğitimi, Gazi Üniversitesi, Bilişim Enstitüsü, 2011.
- [74] Öztemel, E., Yapay Sinir Ağları, İstanbul: Papatya Yayıncılık Eğitim, 2003.
- [75] Alpaydın, E, Introduction to Machine Learning, The MIT Press, 3-6, 2004.
- [76] Dalyan, Tuğba, Makine Öğrenmesinde 1R Algoritması ve İkinci Kuralın (2R) Oluşturulması, Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı, 2006.
- [77] Carbonell, J. G., Michalski, R. S., Mitchell, T.M., Machine Learning: a Historical and Methodological Analysis, AI magazin, 400-408, 1989.
- [78] Nilsson, N., Introduction to Machine Learning, Draft, 1996.
- [79] Han, J., Kamber , M., Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, Massachusetts, USA, 978-0-12-381479-1, 2001.
- [80] Bao, Y., Tsuchiya, E., Ishii, N., Classification by Instance-Based Learning Algorithm, Lecture Notes in Computer Science (LNCS), 3578, 133-140, 2005.
- [81] TM, Mitchell, Machine Learning, McGraw–Hill Companies Press, 1997.
- [82] İşler, Yalçın, A Detailed Analysis of the Effects of Various Combinations of Heart Rate Variability Indices in Congestive Heart Failure, İzmir: Doktora Tezi, Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü.
- [83] AS, Nixon MS ve Aguado, Feature Extraction and Image Processing, Newness Butterworth-Heinemann, Woburn, 1st ed, 2002.
- [84] Öztürk, Uğur, Konjestif Kalp Yetmezliği Hastalarının Teşhisinde Veri Azaltma Yöntemleri Kullanılarak En Yakın Komsu Sınıflandırıcı Çalışma Süresinin Kısaltılması, Zonguldak: Yüksek Lisans Tezi, Bülent Ecevit Üniversitesi, Fen Bilimleri Enstitüsü, 2012.
- [85] S. Vatanabe, Pattern Recognition: Human and Mechanical, New York: JohnWiley ve Sons Inc., 1985.
- [86] Duda R O, Hart P E and Stork D G., Pattern Classification, New York: John Wileyand Sons, 2001.
- [87] Sahan S, Polat K, Kodaz H ve Güneş S, A New Hybrid Method Based On Fuzzy–Artificial İmmune System and K–NN Algorithm For Breast Cancer Diagnosis, Computers in Biology and Medicine, 37:415–423, 2007.
- [88] M, İşler Y ve Kuntalp, Combining Classical HRV İndices with Wavelet Entropy Measures İmproves to Performance in Diagnosing Congestive Heart Failure, Computers in Biology and Medicine, 37(10):1502–1510, 2007a.
- [89] M, işler Y ve Kuntalp, İvestigating Effects of Wavelet Entropy Detailed Measures in Heart Rate Variability Analysis, IEEE 15th Signal Processing and Communications Applications Conference, Turkey, Eskişehir, 2007b.
- [90] M, İşler Y ve Kuntalp, Heart Rate Normalization in The Analysis of Heart Rate Variability in Congesti and Heart Failure., Proceedings of The Institution of

- Mechanical Engineers, Part H: Journal of Engineering in Medicine, 224(3):453-463, 2010.
- [91] Hudairy, H., Data Mining and Decision Making Support in the Governmental Sector, Kentucky: Master Thesis, Louisville University, 1-5, 2004.
- [92] Duda, R. O., Hart, P. E., Stark. Pattern Classification (2nd ed.), John Wiley & Sons, 20-25, 2001.
- [93] Stephen, M. S., Thomas Bayes' Bayesian Inference, Journal of the Royal Statistical Society, Series A, 250–258, 1982.
- [94] Domingos P., Pazzani M., On the Optimality of the Simple Bayesian Classifier under Zero-One Loss, Machine Learning, 29, 103-130, 1997.
- [95] Burhan Yumak, Elektronik Postaların Ayrıştırılmasında Naive Bayesian ve Bulanık Mantık Yöntemlerinin Karşılaştırılması, Ankara: Yüksek Lisans Tezi, Bilgisayar Eğitimi, Gazi Üniversitesi, Bilişim Enstitüsü, 2011.
- [96] Büyükyılmaz, Selçuk, Biyomedikal İşaretlerin Sınıflandırılması, Yüksek Lisans Tezi, Elektronik ve Haberleşme Mühendisliği Anabilim Dalı, 2012.
- [97] Sezer, M., Yapay Sinir Ağları (YSA) Kullanılarak KO Parametresinden BO Parametresinin Tahmin Edilmesi, Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, 2007.
- [98] Kara, Yunus, Statik Senkron Kompanzatorün (STATCOM) Yapay Sinir Ağları (YSA) ile Denetimi, Yüksek Lisans Tezi, Fırat Üniversitesi, Fen Bilimler Enstitüsü, Elektrik Eğitimi Anabilim Dalı, 2008.
- [99] Vemuri V.R., Artificial Neural Networks: Concepts and Control Applications, California: IEEE Computer Society Press, Los Alamitos, p. 42., 1992.
- [100] Trippi R. R. and Turban E., Neural Network in Finance and Investing, Chicago: Irwin Professional Pub., p.4, 1996.
- [101] Yazıcı, Gül, Genetik Algoritmalar ile Radyal Temelli Fonksiyon Ağlarının Optimizasyonu, İstanbul: Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Elektronik ve Haberleşme Anabilim Dalı, 2006.
- [102] H. Yurtoğlu, Yapay Sinir Ağları Metodolojisi ile Öngörü Modellemesi, Yayın No:DPT:2683, Ekonomik Modeller ve Stratejik Araştırmalar Genel Müdürlüğü, 2005.
- [103] T. Kohonen, Self-Organizing Maps, Berlin: Springer, 1995.
- [104] Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., Golub, T.R., Interpreting Patterns of Gene Expression with Self-Organizing Maps : Methods and Application to Hematopoietic Differentiation, Proc. Natl. Acad Sci USA Vol. 96, pages 2907 2912, 1999.
- [105] Özdoğan, Alper, Gen Kümeleme İşleminin Özdüzenleyici Haritalar Kullanılarak Gen Ekspresyonu, Motif Sıklık ve Gen Konum Verilerinden Faydalanılarak Gerçekleştirimi, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, 2009.
- [106] Fausett, Laurene, Fundamentals of Neural Networks: Architectures Algorithms and Applications, Prentice Hall, 1993.
- [107] Manual, TheInternational Classification of Sleep Disorders Diagnostic And Coding, Manual, The International Classification of Sleep Disorders Diagnostic

And Coding, 1990.

- [108] <http://www.edfplus.info/>, Erişim Tarihi: 29 Temmuz 2013.
- [109] http://en.wikipedia.org/wiki/European_Data_Format, Erişim Tarihi: 28 Temmuz 2013.
- [110] S. Pourzare, EEG İşaretlerindeki Göz ve Çene Hareketi Artifaktlarının Sınıflandırılması, Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, 2012.
- [111] Özmen, Nurhan Gürsel, Beyin Bilgisayar Arayüzü Tasarımı için Farklı Zihinsel Aktiviteler Esnasında Oluşan EEG İşaretlerinin Analiz Edilmesi ve Sınıflandırılması, Doktora Tezi, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Makina Mühendisliği Anabilim Dalı, 2010.
- [112] Apshankar, K., Waterhouse, M., Zhang, L. J., O'Riordan, D., Sathwani, D., Web Services Business Strategies And Architectures, Chicago: Expert Press, Illinois, USA, 136-137, 2002.
- [113] Erl, T., Service-Oriented Architecture a Field Guide to Integrating XML and Web Services, Prentice Hall, USA, 2-4, 18-44, 2004.
- [114] Yürekten, Özgür, Hiyerarşik Verilerin XLM Veritabanı Olarak Modellenmesi ve Aralarındaki Benzerliğin Bulunması, Yüksek Lisans Tezi, Bilgisayar Mühendisliği Gazi Üniversitesi, Fen Bilimleri Enstitüsü, 2007.
- [115] Reutemann, David Scuse ve Peter, WEKA Experimenter Tutorial for Version 3-4, Yeni Zelanda Waikato Üniversitesi, , s.1-6, 2007.
- [116] Tekerek, Adem, Veri Madenciliği Süreçleri ve Açık Kaynak Kodlu Veri Madenciliği Araçları, Ankara: Gazi Üniversitesi, Elektronik-Bilgisayar Eğitimi Bölümü.
- [117] Dener, M., Dörterler, M., Orman, A., Açık Kaynak Kodlu Veri Madenciliği Programları: Weka'da Örnek Uygulama, Akademik Bilişim 09 - XI Akademik Bilişim Konferansı Bildirileri, Harran Üniversitesi, Şanlıurfa, 11-13 Şubat 2009.
- [118] Ayık Y.Z., Özdemir A., Yavuz U., Lise Türü ve Lise Mezuniyet Başarısının Kazanılan Fakülte ile İlişkisinin Veri Madenciliği Tekniği ile Analizi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü Dergisi, Cilt 10, S. 2, 441-454, 2007.
- [119] Terzi Ö., Küçüksille E. U., Ergin G., İlker A., Veri Madenciliği Süreci Kullanılarak Güneş Işınımı Tahmini, SDU International Technologic Science, 2011, 3, 29-37.
- [120] Garcia, S., A Study of Statistical Techniques and Performance Measures for Genetics-Based Machine Learning: Accuracy And Interpretability, Soft Computing, 13, 959-977, 2009.
- [121] Sokolova, M. and Lapalme, G., A Systematic Analysis of Performance Measures for Classification Tasks, Information Processing and Management, 45, 427-437, 2009.
- [122] Zanifa Omary, Fredrick Mtenzi, Machine Learning Approach to Identifying the Dataset Threshold for the Performance Estimators in Supervised Learning, Dublin Institute of Technology, Ireland.
- [123] Ferri, C., and Hern, J., An Experimental Comparison of Performance Measures for Classification, Pattern Recogn. Lett., 30, 27-38, 2009.

- [124] Dem, J., Statistical Comparisons of Classifiers over Multiple Data Sets, Journal of Mach. Learn. Res., 7, 1-30, 2006.
- [125] Esmâ Sezer, Epilepsi Teşhisi için EEG Sinyal Analizi, Yüksek Lisans Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Elektronik Bilgisayar Sistemleri, Eğitimi Anabilim Dalı, Konya, 2008.
- [126] Metz CE., Basic Principles of ROC Analysis, Sem Nuc Med., 283-298, 1978.
- [127] Uğurman, F., Berktaş, B., Tıpta Tanı Testleri ve Risk Değerlendirmelerinde Kullanılan Temel Kavramlar, Akciğer Arşivi, 2003.

ÖZGEÇMİŞ

1979 yılında Edirne’de doğdu. İlk ve orta öğrenimini Edirne’de tamamladıktan sonra 2002 yılında Dumlupınar Üniversitesi Elektrik-Elektronik Mühendisliği bölümünden mezun oldu. 2004 yılı başında Trakya Üniversitesi İpsala Meslek Yüksekokuluna öğretim görevlisi olarak atandı. 2009 yılında Trakya Üniversitesi Bilgisayar Mühendisliği bölümü yüksek lisans programından mezun oldu.

ESERLER

A. Ulusal Bilimsel Toplantılarda Sunulan ve Bildiri Kitaplarında Basılan

Bildiriler:

Uyku Apnesi ve EKG Arasındaki İlişkinin Analizi, Tıp Teknolojileri Ulusal Kongresi 2011, (TIPTEKNO'11) İlhan UMUT, **Güven Çentik**, Erdem UÇAR, Levent ÖZTÜRK, Antalya, 13-16 Ekim 2011

Uykuda Periyodik Bacak Hareketi Bozukluğu Hastalığının Uyku Safhalarına Göre Sıklık Analizi, **Güven Çentik**, İlhan Umut, Erdem Uçar, Levent Öztürk Tıp Teknolojileri Ulusal Kongresi 2012 (TIPTEKNO'12), Antalya,1-3 Kasım 2012 sayfa 118-120

Veri Madenciliği Yöntemleri Kullanılarak EEG Sinyallerindeki Epileptiform Aktivitenin Sınıflandırılması, İlhan Umut, Ümit Can Kumdereli, **Güven Çentik**, Erdem Uçar Tıp Teknolojileri Ulusal Kongresi 2012 (TIPTEKNO'12), Antalya, 1-3 Kasım 2012, sayfa 121-124

Görme Engellilere Kabartma Yazı Alfabetini Öğretmek Amacıyla Bilgisayar Destekli Bir Eğitim Setinin Tasarımı ve Uygulaması, **Güven Çentik**, Özlem Uçar, Akademik Bilişim 2013(AB2013), Akdeniz Üniversitesi, Antalya, 23-25 Ocak 2013