

**KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**ELEKTRONİK VE BİLGİSAYAR EĞİTİMİ
ANABİLİM DALI**

YÜKSEK LİSANS TEZİ

**UYKU EVRELEMESİNDE ÇEŞİTLİ DALGACIK VE
SINIFLANDIRICILARIN PERFORMANS ANALİZİ**

ERDEM TUNCER

KOCAELİ 2015

KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRONİK VE BİLGİSAYAR EĞİTİMİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

UYKU EVRELEMESİNDE ÇEŞİTLİ DALGACIK VE
SINIFLANDIRICILARIN PERFORMANS ANALİZİ

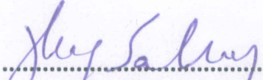
ERDEM TUNCER


Doç.Dr. Emine DOĞRU BOLAT
Danışman, Kocaeli Üniv.

Doç.Dr. Deniz ŞAHİN
Jüri Üyesi, Kocaeli Üniv.

Yrd.Doç.Dr. Seda POSTALCIOĞLU
Jüri Üyesi, Abant İzzet Baysal Üniv.


.....


.....


.....

Tezin Savunulduğu Tarih: 05.06.2015

ÖNSÖZ ve TEŞEKKÜR

Bu çalışmanın hazırlanması esnasında bana yol gösteren, bu alanda çalışmam için yardımlarını ve desteklerini esirgemeyen değerli hocam Doç. Dr. Emine Doğru BOLAT ve fikirleriyle katkıda bulunan Yrd. Doç. Dr. Ömer ŞAYLI'ye çok teşekkür ederim.

Çalışmalarım sırasında değerli katkılarıyla bana yardım eden arkadaşım Oğuzhan TİMÜŞ'a çok teşekkür ederim.

Tez çalışmam sırasında bana her zaman destek olan ve sabır gösteren değerli annem, babam, ağabeyim ve tabi ki eşim Çiğdem TUNCER'e tüm kalbimle teşekkür ediyorum.

Mayıs - 2015

Erdem TUNCER

İÇİNDEKİLER

ÖNSÖZ ve TEŞEKKÜR.....	i
İÇİNDEKİLER	ii
ŞEKİLLER DİZİNİ.....	iv
TABLolar DİZİNİ	vi
SİMGELER ve KISALTMALAR DİZİNİ	vii
ÖZET.....	ix
ABSTRACT.....	x
GİRİŞ	1
1.GENEL BİLGİLER	3
1.1. Uyku Evreleme	4
1.2. EEG İşaretinin Ölçülmesi.....	6
1.3. EOG İşaretinin Ölçülmesi	9
1.4. EEG İşaretlerinin Elektriksel Özellikleri.....	9
1.4.1. Alfa dalgaları	10
1.4.2. Beta dalgaları	10
1.4.3. Teta dalgalar	10
1.4.4. Delta dalgaları.....	10
1.5. Uyku Evreleri	10
1.5.1. Uyanık (WAKE).....	10
1.5.2. Evre 1 (NREM1).....	11
1.5.3. Evre 2 (NREM2).....	11
1.5.4. Evre 3 (NREM3).....	11
1.5.5. Evre R (REM).....	12
1.6. Uyku Evrelerinin Uyku Döngüsü Boyunca Dağılımı	12
2.MATERYAL ve YÖNTEM.....	13
2.1. Özellik Çıkartımı	13
2.1.1. Dalgacık dönüşümü	13
2.1.1.1. Sürekli dalgacık dönüşümü	16
2.1.1.2. Ayırık dalgacık dönüşümü	18
2.1.1.3. Dalgacık paket dönüşümü	20
2.1.1.4. Dalgacık dönüşümünün kullanım alanları	21
2.1.2. Enerji.....	21
2.1.3. Maksimum tepe değeri.....	22
2.1.4. Minimum tepe değeri.....	22
2.1.5. Standart sapma	23
2.1.6. Ortalama değer.....	24
2.1.7. Entropi	24
2.1.8. Hjorth parametreleri.....	25
2.2. Makine Öğrenmesi Teknikleri.....	26
2.2.1. Özellik azaltımı (Seçimi)	26
2.3. Veri Madenciliği.....	29
2.3.1. Birliktelik kuralı.....	30
2.3.2. Kümeleme modelleri (Clustering)	31

2.3.3. Sınıflandırma (Classification) ve regresyon (Regression) modelleri	31
2.3.3.1. KNN komşuluğu (KNN neighbours)	32
2.3.3.2. Destek vektör makinası (Support vector machine)	33
2.3.3.3. LR (Lojistik regresyon) modeli.....	33
2.3.4. Değerlendirme ölçütleri	39
2.3.4.1. Sınıflandırma doğruluğu	39
2.3.4.2. Karmaşıklık matrisi (Confusion matrix)	39
2.4. Uygulamada Kullanılan Veri Madenciliği Yazılımı	43
2.4.1. Weka	43
2.4.1.1. Weka da kullanılan özellikler	48
2.4.1.2. Weka dosya formatı	48
2.4.1.3. Weka sonuç ekranı	49
3. BULGULAR	51
3.1. EEG - EOG Sinyallerinin Alınması	53
3.2. Ön İşleme	54
3.3. Segmentasyon.....	56
3.4. Öznitelik Çıkartımı.....	57
3.5. Öznitelik Azaltımı	60
3.5.1. Sarmal metot ile KNN sınıflandırıcısı	60
3.5.2. Sarmal metot ile lojistik regresyon	62
3.5.3. Sarmal metot ile SVM sınıflandırıcısı	64
3.5.4. Sarmal metot ile ayırt edici öznitelikler.....	65
3.5.6. Filtre metodu ile KNN sınıflandırıcısı	67
3.5.7. Filtre metodu ile lojistik regresyon sınıflandırıcısı.....	68
3.5.8. Filtre metodu ile SVM sınıflandırıcısı	69
3.5.9. Filtre metoduna göre ayırt edici öznitelikler.....	73
4. SONUÇLAR ve ÖNERİLER.....	75
KAYNAKLAR	77
KİŞİSEL YAYIN ve ESERLER.....	81
ÖZGEÇMİŞ	82

ŞEKİLLER DİZİNİ

Şekil 1.1.	Tüm gece uykusunun uyku dönemlerine dağılımı.....	5
Şekil 1.2.	Uluslararası 10-20 sistemine göre, 19 kanallı elektrot yerleşimi.....	7
Şekil 1.3.	EEG monopolar ölçüm tekniği	8
Şekil 1.4.	EEG bipolar ölçüm tekniği	8
Şekil 1.5.	CZA1 kanalından alınmış uyanık evresine ait örnek 30 sn'lik epok.....	11
Şekil 1.6.	Uyku evrelerinin 8 saatlik uyku dönemindeki döngüsü	12
Şekil 2.1.	Dalgacığın zamanda ötelenmesi	14
Şekil 2.2.	Dalgacığın ölçeklenmesi	14
Şekil 2.3.	Dalgacık çeşitleri	15
Şekil 2.4.	Zamana göre frekansı değişen sinüs dalgası.....	15
Şekil 2.5.	Zamana göre frekansı değişen sinüs dalgasının fourier dönüşümü	16
Şekil 2.6.	Zamana göre frekansı değişen sinüs dalgasının dalgacık spektrumu.....	16
Şekil 2.7.	Sürekli dalgacık dönüşümü.....	17
Şekil 2.8.	Bir zaman serisinin aşağı örneklenme algoritması	19
Şekil 2.9.	Gerçek zamanlı serinin aşağı örneklenmesi.....	20
Şekil 2.10.	200 örneklili bir sinyalin 2.seviye ayrışım basamakları	20
Şekil 2.11.	Paket dalgacık dönüşümü 2. seviye ayrışım basamakları.....	21
Şekil 2.12.	Ayrık ve sürekli enerji işaretleri	22
Şekil 2.13.	EEG kaydı	23
Şekil 2.14.	Değerlendirici	28
Şekil 2.15.	Veri Madenciliği	30
Şekil 2.16.	Sınıflandırma modelinde kullanılan başlıca yöntemler	32
Şekil 2.17.	KNN sınıflandırıcısı.....	33
Şekil 2.18.	Destek vektör makinesi.....	35
Şekil 2.19.	Lojistik Regresyon	38
Şekil 2.20.	Weka Genel Ara yüz görünümü	44
Şekil 2.21.	Weka veriler yüklü değil iken ara yüz görünümü.....	44
Şekil 2.22.	Weka veriler yüklü iken elde edilen ara yüz görünümü	45
Şekil 2.23.	Weka özellik azaltımı	46
Şekil 2.24.	Epokların uyku evresi sınıflarına dağılımı.....	47
Şekil 2.25.	Eğitim ve test verisi oranı ekran görünümü.....	47
Şekil 2.26.	Weka genel dosya formatı	48
Şekil 2.27.	Çalışmada kullanılan weka dosya formatı	49
Şekil 2.28.	Weka sonuç ekranı	50
Şekil 3.1.	Önerilen sistemin akış diyagramı.....	53

Şekil 3.2.	Edf uzantılı verinin Matlab ortamına aktarılması	54
Şekil 3.3.	Filtre Cevabı	55
Şekil 3.4.	Standart sapmanın sabitlenmesiyle sınıflandırıcı başarısının değişimi	55
Şekil 3.5.	10 dk'lık EEG kaydı	56
Şekil 3.6.	30 sn'lik EEG kaydı.....	56
Şekil 3.7.	Ayrışım seviyeleri.....	57
Şekil 3.8.	Öznitelik azaltımı yapılmadan dalgacık tipine göre başarı oranları	58
Şekil 3.9.	Dalgacık tipine göre KNN sınıflandırıcısı başarısı.....	61
Şekil 3.10.	Dalgacık tipine göre KNN sınıflandırıcı için seçilen öznitelik sayısı.....	62
Şekil 3.11.	Dalgacık tipine göre lojistik regresyon sınıflandırıcısı başarısı.....	63
Şekil 3.12.	Dalgacık tipine göre lojistik regresyon sınıflandırıcı için seçilen öznitelik sayısı.....	64
Şekil 3.13.	Dalgacık tipine göre destek vektör makinası sınıflandırıcısı başarısı	64
Şekil 3.14.	Dalgacık tipine göre destek vektör makinası sınıflandırıcı için seçilen öznitelik sayısı	65
Şekil 3.15.	Sarmal metoda göre özniteliklerin seçilme oranı	66
Şekil 3.16.	Filtre metodu lojistik regresyon sınıflandırıcısıyla öznitelik azaltım başarı oranları	69
Şekil 3.17.	Filtre metodu ile SVM sınıflandırıcısı başarı oranı	70
Şekil 3.18.	Dalgacık Çeşitlerine göre uyku evresi sınıflandırma başarısı.....	73

TABLULAR DİZİNİ

Tablo 1.1.	AASM standartları ile önerilen teknik ve dijital özellikler.....	6
Tablo 1.2.	EEG sinyali frekans bantları	9
Tablo 2.1.	Örnek öznelik azaltımı tablosu	29
Tablo 2.2.	Karmaşıklık matrisi	40
Tablo 2.3.	Doğru Pozitif.....	41
Tablo 2.4.	Yanlış pozitif.....	42
Tablo 2.5.	Yanlış negatif.....	42
Tablo 2.6.	Doğru negatif	43
Tablo 2.7.	Karmaşıklık Matrisi	43
Tablo 3.1.	Uyku evresi skorlaması yapılan hastalara ait istatistiksel veriler	52
Tablo 3.2.	EEG sinyali alt frekans bantları	58
Tablo 3.3.	Bir epok için elde edilen öznelikler	59
Tablo 3.4.	KNN (5-9-11) sınıflandırıcı öznelik sayısına göre elde edilen başarı oranları.....	68
Tablo 3.5.	Her bir dalgacık için karmaşıklık matrisi	71
Tablo 3.6.	Coif 5 dalgacığı için karmaşıklık matrisinin türevleri	71
Tablo 3.7.	Db 4 dalgacığı için karmaşıklık matrisinin türevleri	72
Tablo 3.8.	Db 44 dalgacığı için karmaşıklık matrisinin türevleri	72
Tablo 3.9.	Sym 8 dalgacığı için karmaşıklık matrisinin türevleri.....	72
Tablo 3.10.	Uyku evresi tespitinde başarı oranı yüksek olan öznelikler	74

SİMGELER ve KISALTMALAR DİZİNİ

Dk	: Dakika
μ V	: Mikrovolt
Hz	: Herz
Sn	: Saniye

Kısaltmalar

A	: Approximate (Yaklaşık)
AASM	: American Academy of Sleep Medicine (Amerikan Tıbbi Uyku Akademisi)
ADD	: Ayrık Dalgacık Dönüşümü
ARFF	: Attribute Relationship File Format (İlişkili Dosya Biçimi Özelliği)
C	: Central (Merkez)
COIF	: Coiflets
D	: Detail (Detay)
DB	: Daubechies
EEG	: Elektroensefalogram
EKG	: Elektrokardiyografi
EMG	: Elektromiyografi
EOG	: Electrooculography (Elektrookülografi)
F	: Frontal (Ön)
FC	: Frontal Cental (Ön Merkez)
FP	: Frontal Polar (Ön Kutup)
IIR	: Infinite Impulse Response (Sonsuz birim Dürtü Cevabı)
KNN	: K-Nearest Neighbors (En Yakın K Komşuluğu)
LR	: Logistic Regression (Lojistik Regresyon)
ML	: Maksimum Likelihood (Maksimum Benzerlik)
NREM	: Non Rapid Eye Movement (Hızlı Olmayan Göz Hareketleri)
O	: Oksipital
P	: Pariyetal
PO	: Pariyetal – Oksipital
P-P	: Peak to Peak (Tepeden Tepeye)
PSG	: Polisomnografi
PUK	: Pearson VII Universal
RBF	: Radial Basis Function (Radyal Temelli Fonksiyon)
REM	: Rapid Eye Movement (Hızlı Göz Hareketleri)
R&K	: Rechtschaffen ve Kales
S	: Ölçek
SDD	: Sürekli Dalgacık Dönüşümü
SEM	: Slow Eye Movements (Yavaş Göz Hareketleri)
SYM	: Symlet
SVM	: Support Vector Machine (Destek Vektör Makinası)
τ	: Öteleme

T : Temporal (Şakak)
TXT : Text (Metin)
VAR : Varyans
WEKA : Waikato Bilgi Analizi Ortamı

UYKU EVRELEMESİNDE ÇEŞİTLİ DALGACIK VE SINIFLANDIRICILARIN PERFORMANS ANALİZİ

ÖZET

Bu çalışmada, uyku evrelerinin sınıflandırılması için EEG (Elektroensefalogram) ve EOG (Electrooculography) işaretlerinin dalgacık analizi ile incelenip, KNN (K-nearest neighbors) komşuluğu, destek vektör makinaları ve lojistik regresyon sınıflandırıcıları kullanılarak en yüksek başarı oranının elde edilmesi amaçlanmıştır. Alınan EEG işaretleri Matlab programında farklı dalgacık tipleri ile dalgacık dönüşümü kullanılarak alt frekans bantlarına ayrılmış ve bu alt bantlardan her bir epok için öznelikler elde edilmiştir. Çıkarılan öznelik boyutlarının azaltımı için Weka (Waikato Environment for Knowledge Analysis) programı aracılığıyla farklı öğrenme algoritmaları olan sarmal (Wrapper) ve filtre metotları kullanılmış ve sınıflandırıcılara uygulanarak başarı oranları elde edilmiştir. Kullanılan dalgacık çeşitlerinin ve sınıflandırıcıların performanslarını karşılaştırmak amacıyla karmaşıklık matrisi analizleri kullanılmıştır.

Anahtar Kelimeler: Dalgacık Dönüşümü, Elektroensefalografi(EEG), Sinyal İşleme, Uyku Evreleri Sınıflandırılması, Weka.

PERFORMANCE ANALYSIS OF VARIOUS WAVELETS AND CLASSIFIERS FOR SLEEP STAGING

ABSTRACT

In this study, EEG (Elecktroensefalogram) and EOG (Electrooculography) based sleep stage classification is realised using wavelet analysis, KNN (K-nearest neighbors), support vector machine and logistic regression for getting the highest performance. The taken EEG signals are divided into the sub bands employing different wavelet types and wavelet transform in Matlab program. The features for each epoch are derived from the sub bands. Feature selection process is realized by Weka (Waikato Environment for Knowledge Analysis) program utilizing wrapper and filter methods and the selected features are applied to the classifiers and the performance rates are achieved. Confusion matrix analysis is employed for comparing the used wavelet types and the classifier performances.

Key Words: Wavelet Transform, Elektroensefalografi (EEG), Signal Processing, Sleep Stage Classification, Weka.

GİRİŞ

Her insan, ömrünün üçte birini uyuyarak geçirir. Bu oran, bebeklik, çocukluk ve gençlik dönemlerinde artmakta, yetişkinlik ve ileri yaşlarda ise azalmaktadır. Uyku vücut sisteminin pasif hale geldiği bir dönem değildir. Aksine vücudun aktif bir yenilenme sürecine girdiği dönemdir. Uyku kişinin sağlığı için çok önemlidir. Uyku kalitesi insan hayatını doğrudan etkileyen en önemli unsurlardan bir tanesidir. Uyku eksikliği beynin çalışmasını ciddi ölçüde etkiler. Unutkanlık, dalgınlık, kendini rahatsız hissetmek gibi durumlar yaşanabilir. Bu nedenle hekimlerin uyku kalitesini değerlendirebilmesi için bir araç gereklidir. PSG(Polisomnografi) cihazı ile temelde uykunun yapısı ve fizyopatolojik değişimler araştırılır. Uykunun yapısını, uykudaki psikolojik, biyolojik ve patolojik değişimleri uyku evreleri ile olan ilişki içerisinde ortaya koyar. Tıp alanında hastalıkların tanılarında ve bu alanda yapılan çalışmalarda kullanılan EEG (Elektroensefalogram) işaretleri beyinde oluşan sinirsel aktivitelerin sonucunda elde edilen elektriksel işaretlerdir. Bu işaretler beynin fonksiyonları ve çalışması ile ilgili çok miktarda bilgi barındırır. Bu bilgilerin kullanılabilir hale gelmesi için sinyal işleme alanındaki uzmanlar çok çeşitli araştırmalar yapmaktadırlar. Uyku sırasında, insan beyni birbirleriyle ilişkili evrelerden geçer. Çoğu sinir merkezi pasif duruma geçer ve daha az karmaşık hale gelir. Beyin modellenmek için daha elverişli duruma gelir [1]. EEG beyin tarafından üretilen elektriksel aktivitedir. Beynin aktivitesini ölçerken tek bir sinir hücresi aktivitesini ölçmek mümkün değildir. EEG işareti birçok sinir hücresi etkileşiminin elektriksel sonucudur ve insan derisi yüzeyine bağlanan elektrotların bağlanmasıyla ölçülür. Ölçüm sonucu elde edilen işaret, çok küçük genlikli gerilim değişiklikleridir. Uykuyu gece boyunca birkaç aşamadan oluşan döngü oluşturur. Vücut uykuya hazırlanırken, kaslar gevşer ve yavaş göz hareketleri belirir [2].

EEG, EKG(Elektrokardiyografi) ve EMG (Elektromiyografi) işaretlerine benzer şekilde frekans bantlarına göre değerlendirilmektedir. Çünkü EEG işaretleri periyodik değildir. Genlik, faz ve frekanslar sürekli olarak zamana göre değişir. EEG işaretlerinin temel frekansı ile beyin aktivitesi yakından ilişkilidir.

Geleneksel EEG sinyali ile PSG kullanılarak izlenir. Ölçümler genellikle uyku laboratuvarında alınır. Manuel uyku puanlama zaman alıcı ve hata payı yüksek bir süreçtir. Bu nedenle, klinik ve deneysel çalışmalarda kullanılacak kolayca uygulanabilir yöntemlere ihtiyaç vardır [3].

Bu tez çalışmasında, elde edilen EEG ve EOG sinyallerinden uyku evrelerinin tespiti için dalgacık dönüşümü kullanılmıştır. Her bir banttaki frekans aralıkları, sinyal işleme teknikleriyle elde edilmiş öznitelikler kullanılarak, sınıflandırma yöntemleri ile başarı oranları tespit edilmiştir. AASM (American Academy of sleep medicine) standartlarına göre uyku evrelerinin doğru şekilde tespit edilmesi amaçlanmıştır.

Birinci bölümde, tez çalışmasına giriş yapılmış uyku ve uyku evreleri hakkında genel bilgilere yer verilmiş, ikinci bölümde, bu tez çalışmasında kullanılan materyal ve yöntemlerden bahsedilmiştir. Üçüncü bölümde, kullanılan yöntem ve teknikler ile elde edilen bulgular anlatılmış aynı zamanda yapılan çalışmalar hakkında tanımlamalar yapılmıştır. Dördüncü bölümde, çalışmalar neticesinde ortaya çıkan sonuçlar sunulmuş, karşılaştırılmış ve önerilerde bulunulmuştur.

Son olarak beşinci bölümde, faydalanılan kaynaklar verilmiştir.

1. GENEL BİLGİLER

Uyku; canlılarda enerjinin korunmasını, sinir sisteminin gelişim ve onarımını sağlayan süreç olup; uyarılmışlığı, bilişsel işlevleri denetleyen biyolojik oluşumun birçoğu ile iç içedir [1-3]. Berger tarafından 1929 yılında başlatılan EEG ile ilgili çalışmalar, uykunun incelenmesinde büyük bir başlangıç olmuştur. Uyku anındaki EEG kayıtları ile ilgili araştırmalar 1937 yılında Loomis ve arkadaşları tarafından yapılmıştır [2]. Bu çalışmaları sonucunda uyku beş evre olarak belirlenmiştir.

Uykunun REM dönemini 1957 yılında Dement ve Kleitman tanımlamışlar ve uykunun periyodik olarak birbiri ardına sıralanan dönemlerden oluştuğunu göstermişlerdir [2].

Birçok uyku hastalığı olmasından dolayı uyku ve uyku kalitesinin tespitinin önemi gün geçtikçe artmaktadır. En sık görülen birkaç hastalık, uykusuzluk, narkoepilepsi, uyku apnesidir. Bunların dışında birçok hastalık kendisini uyku sırasında gösterir. Ağrıdan sonra en sık rastlanan hastalık uyku bozukluklarıdır [4].

Uyku adı verilen periyodik, fizyolojik ve geri dönen bilinç ve davranış değişikliği dönemlerinin ne olduğu, nasıl oluştuğu, o sırada neler yaşandığı ve ne işe yaradığı gibi temel bazı sorular, insanlık tarihi kadar eskidir. Uykunun üzerini örten sır perdesinin aralanabilmesi, EEG sinyalinin Hans Berger tarafından keşfinden sonra olmuştur. Berger ilk defa oğlunda saçlı deriden kortikal bioelektrik aktiviteyi yazdırmış ve kısa bir süre sonra uyuyan insanda kayıtlar yapmaya başlamıştır. Bundan sonra konuya ilişkin bilimsel çalışmalar günümüze dek artarak gelmiştir [1]. Uyku sırasında insan beyni birçok psikofizyolojik evrelerden geçer. Birçok sinir merkezi pasifleşir. Bu sayede beyin daha az karmaşık durumdadır ve sistemsal olarak modellemeye elverişli hale gelir.

1.1. Uyku Evreleme

Uyku evreleme, uyku sırasında alınan biyolojik işaretlerin isimlendirilmesidir. Bu işlem uzman doktor tarafından yapılır ve uzun sürmektedir. Bu işlemlerin otomatik olarak yaptırılması işlem yükünü azaltmakta ve hastalıkların teşhis sürecini kısaltmaktadır [3].

“İlk kez 1968 yılında Rechtschaffen ve Kales’in (R&K) editörlüğünde, insan uyku evrelerinin standart terminoloji, teknik ve skorlama el kitabı hazırlanmış ve yakın zamana kadar skorlamada bu kitaptan yararlanılmıştır. Buna göre sağlıklı bir kişide normal uyku iki ana bölüm ve 5 evreden oluşmakta idi” [5].

NREM (Non Rapid Eye Movement) 4 evreden oluşur;

- NREM1
- NREM2
- NREM3
- NREM4

REM

“1999 yılında AASM , uyku sırasında ortaya çıkan anormal solunumsal olaylar ve bunlarla ilişkili sendromların standart ölçüm yöntemlerini saptamak amacıyla bir toplantı düzenlemiştir” [6].

Bu toplantıdan beri AASM skorlama kriterleri zaman zaman güncellenmektedir. 2007 yılında AASM uyku ve uyku ile ilgili olayların skorlanmasına ilişkin kurallar, terminoloji ve teknik özellikleri içeren bir el kitabı yayınlamıştır. Günümüzde, bu skorlama kriterleri temel alınmaktadır. Bu çalışmada, 2014 yılında “Manual for Scoring Sleep-update to Version 2.1” adı ile yayınlanan AASM kılavuzunda yer alan kriterlere göre uyku evreleri skorlaması ele alınmıştır. Bu kılavuza göre NREM uyku evreleri 3 ve 4 olarak değil, sadece 3 olarak skorlanmaktadır. Uyku ise toplamda 4 evreye ayrılmaktadır:

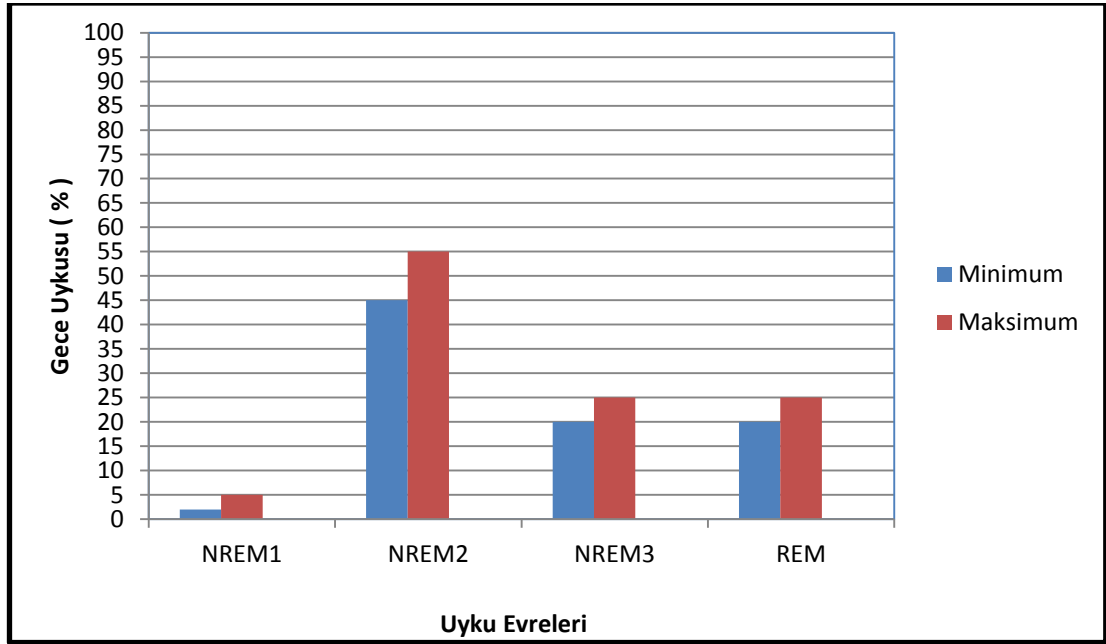
NREM;

- NREM1
- NREM2
- NREM3

REM

Genç erişkinlerde ilk uykuya dalındığında yarı şuurlu olan bir süreçten geçilir. Bu süreç sonunda kısa bir zaman için rüyalı uykuya ardından da derin ve rüyasız bir uykuya girilir. Bu anda beyin dalgası frekansları yavaşlar. Bu evre NREM3 evresi olarak adlandırılmaktadır. Uykuya daldıktan yaklaşık 90 dk sonra uykunun en hafif dönemine ulaşılır. Bu dönemde, uyanma ihtimali çok yüksektir ve çevremizdeki seslere, hareketlere duyarlı hale gelinir. NREM 1,2,3 evreleri ve ardından ortaya çıkan REM uykusu evresine “uyku siklusu” denir. Bu noktada tekrar uykuya dönülürse bu devirler yeniden başlayacaktır. Gece boyunca bu devirler devam ederek, periyodik bir şekilde, derin uyku, rüya görme ve hafif uyku döngüsü devam edecektir. Uyku devrinin tam ortasında uyanılırsa, halsiz ve yorgun hissetme ihtimali artar [7].

Gece uykusunun %2 - %5’lik kısmı NREM1, %45 - %55’lik kısmı NREM2, %20 - %25’lik kısmı NREM3, %20 - %25’lik kısmının ise REM döneminin oluşturduğu Şekil 1.1’de görülmektedir [4-6].



Şekil 1.1. Tüm gece uykusunun uyku dönemlerine dağılımı

Klasik olarak PSG kayıtlarının skorlanması deneyimli bir uyku teknisyeni tarafından elle ve görsel olarak yapılır. Uyku evreleri skorlanırken kaydın tümü sayfalara bölünür, bu sayfaların her birine bir epok adı verilir. PSG kayıtlarından alınan her bir epok deneyimli uyku teknisyeni tarafından monitörden takip edilerek skorlanır [6].

Bir epok AASM standartlarına göre en fazla 30 sn olmalıdır. Son yıllarda geliştirilen yazılım programları sayesinde, teknisyen desteği olmadan, daha kısa sürelerde otomatik skorlama yapılabilmektedir [8].

- “AASM’ye göre uyku evreleri skorlanırken aşağıdaki kurallar geçerlidir:
1. Uykunun skorlanması, uykunun başlangıcından itibaren 30 saniyelik epoklarla yapılmalıdır.
 2. Her epok ayrı ayrı evrelendirilmelidir.
 3. Bir epokta iki ya da daha fazla uyku evresi varsa, epogun çoğunluğunu oluşturan evre skorlanmalıdır” [9].

Uyku evresi tespitinde AASM standartları ile önerilen bazı teknik ve dijital özellikler Tablo 1.1’de gösterilmiştir.

Tablo 1.1. AASM standartları ile önerilen teknik ve dijital özellikler [10]

Elektrot Empedansı (maks)	5 K Ω
Örnekleme Frekansı EEG (min)	200 Hz
Örnekleme Frekansı EOG (min)	200 Hz
Örnekleme Frekansı EKG(min)	200 Hz
Örnekleme Frekansı EMG(min)	200 Hz
Filtre EEG	0,3-35 Hz
Filtre EOG	0,3-35 Hz

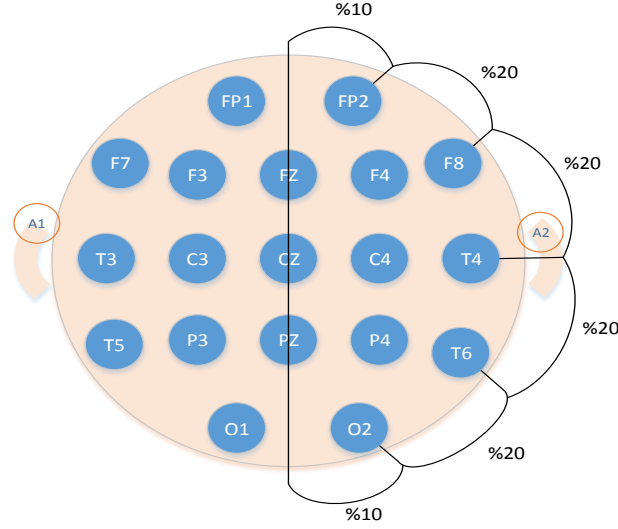
1.2. EEG İşaretinin Ölçülmesi

EEG dalgalarının ölçülmesi ve elektrotların kafa derisine yerleştirilmesi için Uluslararası EEG Federasyon Birliği (International Federation of EEG Societies) tarafından belirlenmiş olan bir standart kullanılır.

EEG'de Uluslararası 10-20 sistemine göre en az 19 kanallı EEG kullanılır. Uluslararası 10-20 sistemi ile elektrotları yerleştirmek, tüm saçlı derinin aynı şekilde kaplanmasını sağlar. Bu sistem baş üzerindeki kemik işaret noktaları arasındaki

uzaklıkları kullanır, başı boydan boy kat eden ve toplam uzunluklarının %10 veya %20'si kadar olan aralıklarla bölünen bir çizgi sistemi yaratır [11].

Tek sayılı elektrotlar sol yarı küreye yerleştirilirken, çift sayılı elektrotlar sağ yarı küreye yerleştirilmiştir.



Şekil 1.2. Uluslararası 10-20 sistemine göre, 19 kanallı elektrot yerleşimi [11]

10/20 sistemine göre elektrot yerleşimlerinin isimlendirilmesi, elektrot ismi harfine eşlik eden bir rakamdan oluşur. Harfler elektrotun üzerinde bulunduğu beyin loblarını ifade eder.

Buna göre:

Fp: Frontal kutup (pole)

F: Frontal

C: Santral (central)

T: Temporal

P: Pariyetal

O: Oksipital

İki harfin kombinasyonu ise ara elektrot yerleşimlerini gösterir, örneğin;

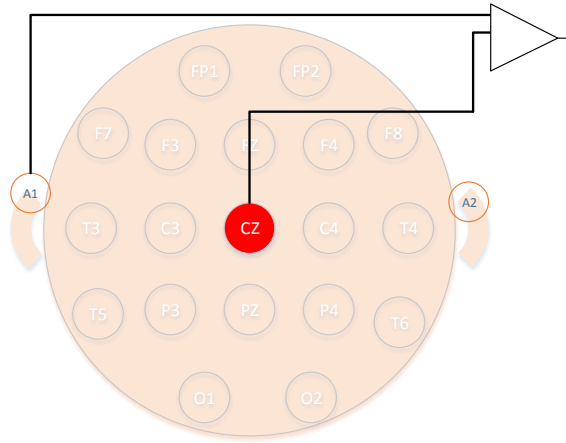
FC: frontal ve santral elektrot yerleşimlerinin arasında yer alır.

PO: pariyetal ve oksipital elektrot yerleşimlerinin arasında yer alır.

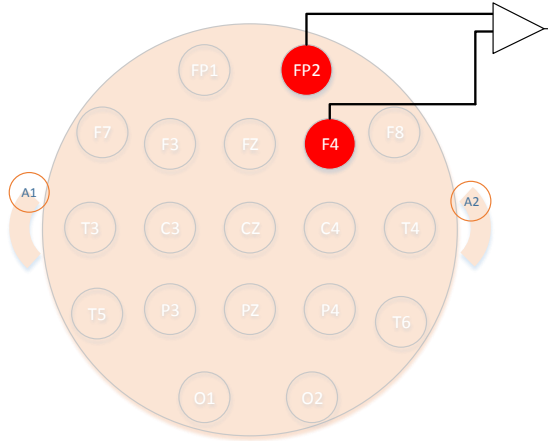
Kafanın sol tarafındaki elektrotlar tek sayılar ile, sağ tarafındaki elektrotlar çift sayılar ile, orta hat elektrotları ise küçük "z" harfi ile gösterilir.

İki adet EEG sinyali ölçüm modu vardır. Bunlar monopolar ve bipolar modlardır;

Monopolar teknikte, potansiyeller bir referans elektrotuna göre ölçülür. Peş peşe bağlanan elektrotlar arasındaki potansiyel farklar ise bipolar kayıt tekniği ile ölçülür [12].



Şekil 1.3. EEG monopolar ölçüm tekniği [12]



Şekil 1.4. EEG bipolar ölçüm tekniği [12]

Uyku evrelerinin skorlanması için en az bir kanal EEG kaydı olmalıdır. Bu kanal, uyku içciklerinin ve K komplekslerinin iyi şekilde gözlemlenebildiği C3 veya C4 bölgeleri olacak şekilde tercih edilmektedir. Bundan dolayı tavsiye edilen derivasyonlar C3A2 ya da C4A1'dir [13].

1.3. EOG İşaretinin Ölçülmesi

EOG bir göz küresi hareketidir. Göz hareketleri medial ve lateral rectus kası, superior ve interior rectu kası, superior ve inferior oblik kası olmak üzere 3 ayrı kas çifti tarafından sağlanır. Gözün iç ve dış köşelerine yerleştirilen elektrotlar bu kasların hareketi vasıtasıyla ölçüm yapar. EOG sinyalleri genliği 50-3500 μ V arasında değişir [12].

“Uyku sırasında göz hareketleri kaydının iki temel nedeni vardır. Birincisi; REM döneminin en karakteristik bulgusu olan hızlı göz hareketleri saptanır. Böylece, uyku evrelemesi için çok önemli bir kriter kaydedilmiş olur. İkincisi ise; uykunun başlangıcında SEM (yavaş göz hareketleri) görülür ve NREM evre-1'e geçişin bulgusudur. Skorlama için kesin kriter olmamakla birlikte oldukça değerli bilgi sağlar. EOG kaydı için en az 2 kanal gereklidir” [13].

1.4. EEG İşaretlerinin Elektriksel Özellikleri

EEG işaretlerinde genlik, faz ve frekans değerleri zaman içerisinde sürekli olarak değişir. EEG işareti Beta, Alpha, Theta ve Delta olmak üzere dört sinyal bandında incelenmektedir. EEG işaretleri, kapsadıkları frekans bantları ve bu bantlara verilen isimler ile birlikte Tablo 1.2'de gösterilmiştir.

Tablo 1.2. EEG sinyali frekans bantları

EEG Sinyal Çeşitleri	Frekans (Hz)
Beta	> 13 Hz
Alpha	8 – 13 Hz
Theta	4 – 8 Hz
Delta	< 4 Hz

1.4.1. Alfa dalgaları

Alfa ritmi uyanık (normal), relaks, istirahatte, sakin kimselerde ve gözler kapalı iken görülür. 8-13 Hz aralığında bulunan dalgalarıdır [11]. Yoğun şekilde oksipital bölgede ortaya çıkar.

1.4.2. Beta dalgaları

Aktif düşünme, ilgilenme, konsantre olabilme, günlük problemleri çözebilme ile ilgili durumda olan kimselerde ve gözler açık iken görülür. Beta dalgaları, frekansları 13 Hz'in üzerindeki beyin dalgalarıdır [11]. Frontal bölgelerde belirgin olarak kaydedilebilir

1.4.3. Teta dalgaları

Teta dalgaları uykuya geçerken veya uykunun ilk evrelerinde olan kimselerde görülür. 4-8 Hz arasındaki dalgalarıdır. Genlikleri 100 μ V (p-p)' den küçüktür [11].

1.4.4. Delta dalgaları

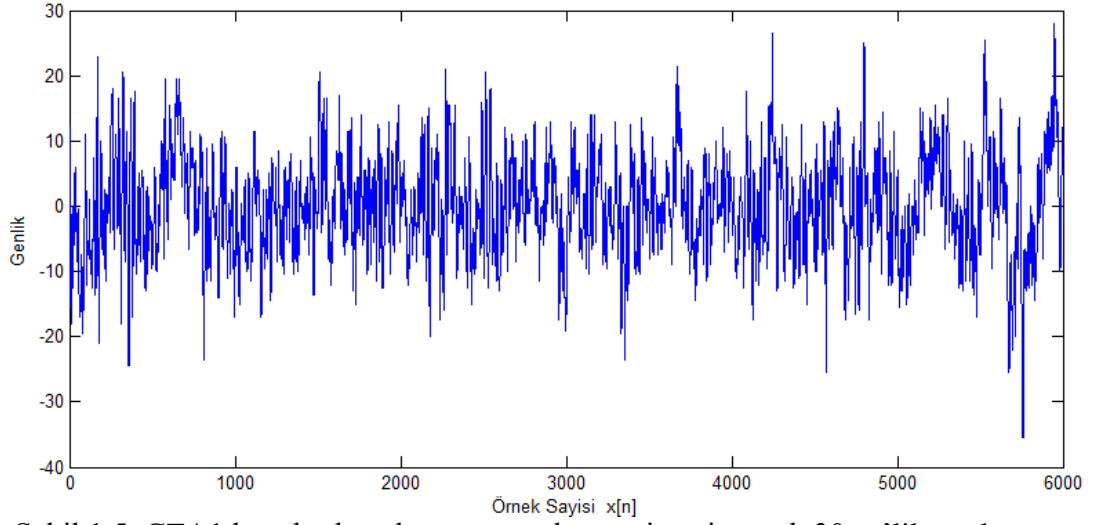
Delta dalgaları derin uyku halindeki kimselerde görülür. Delta dalgaları 0,5–4 Hz aralığındaki beyin dalgalarıdır. Genlikleri 100 μ V (p-p)' den küçüktür. Daha çok frontal bölgelerde görülür [11].

1.5. Uyku Evreleri

Uykuda birbirini periyodik olarak izleyen iki değişim dönemi vardır. Bunlar REM ve NREM dönemleridir. Yüzeysel uykudan derin uykuya gidildikçe dalgaların genlik değerleri artmakta, frekansı azalmaktadır. Uyumak için gözleri kapatmak ile uykuya dalma arasında kalan döneme uykuya dalmanın latent dönemi denir. Bu latent dönemden sonra NREM ve REM uykusu başlar [7].

1.5.1. Uyanık (Wake)

Gözler kapalı iken Okpisital bölgeden alınan alfa dalgaları izlenir. Gözler açılınca bu dalgalar zayıflar ve yerini beta dalgalarına bırakır. EOG kanallarında ise hızlı göz hareketleri belirir.



Şekil 1.5. CZA1 kanalından alınmış uyanık evresine ait örnek 30 sn'lik epok

1.5.2. Evre 1 (NREM1)

Uyku ile uyanıklılık arasında kalınan bir uyku evresidir. Bir kişi NREM1 evresinde uyandırılırsa çevresinde gelişen olayların farkında olmamasına karşın genelde uyanık olduğunu söyler. 4-8 Hz'lik frekans aralığında olan Teta aktivitesi baskın durumdadır. Verteks keskin dalgalar görülebilir. En az 0,5 sn'nin üzerinde oluşan yavaş göz hareketlerinin varlığı da NREM1 evresinin göstergesidir [4,11].

1.5.3. Evre 2 (NREM2)

Uykunun bu evresinde bilinç, kişi uyandırıldığında uykuda olduğunu hatırlayabilmesine yeterli olacak şekilde oluşan bir uyku evresi durumudur.

NREM2 evresinde, negatif keskin ve onu takip eden daha yavaş pozitif komponenti olan K Kompleksi vardır. 12-14 Hz aralığında 0,5 sn'den daha uzun süren peş peşe küçük dalgardan oluşan uyku içcikleri de NREM2 evresinin göstergesidir [4,11].

1.5.4. Evre 3 (NREM3)

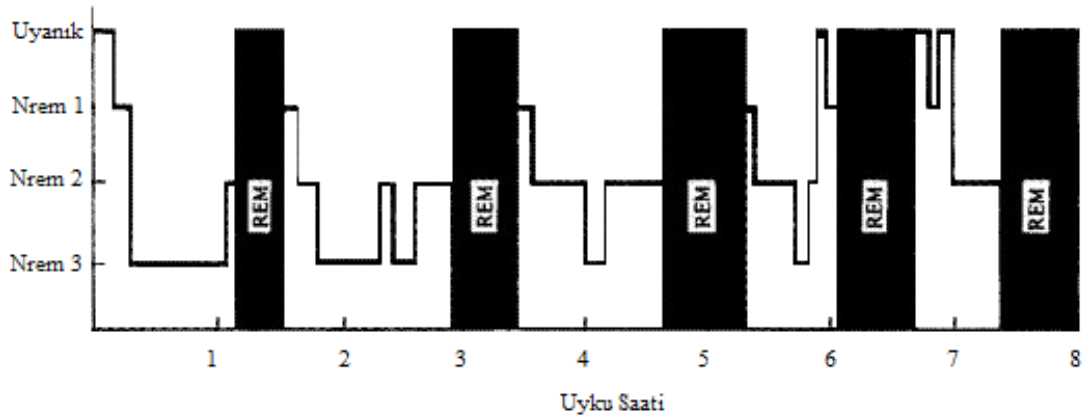
Derin uyku olarak adlandırılan bu evrede delta dalgaları baskındır. Kas ya da göz hareketleri en asgari düzeydedir ve tipik olarak görülmez. En dinlendirici uyku evresidir ve eksik kalması gün içinde uykululuk haline yol açar. Frontal bölgeden alınan ölçümlerde en az $75\mu\text{V}$ genliklidir [4,11].

1.5.5. Evre R (REM)

REM esnasında beynin çalışması uyanıkken ki çalışmasına benzer olduğu için hatırlama olgusu REM uykusu esnasında gerçekleşir. Bu evre uykunun diğer evrelerinin arasına serpiştirilmiştir. Çok sayıda farklı özellik ile bağlantılıdır. Aynı zamanda paradoksal uyku olarak da bilinmektedir. Maksimum 4-8 Hz aralığında keskin uçlu, üçgene benzeyen testere dişi dalgalar ve en az 0,5 sn'nin üzerinde oluşan düzensiz hızlı göz hareketlerinin varlığı da REM evresinin göstergesidir. NREM 1 evresinde olduğu gibi teta aktivitesi baskın durumdadır. Uykunun uyanıklığa en yakın olan evresidir. Bu yüzden çevredeki seslere, hareketlere karşı her an tepki verilebilir [4].

1.6. Uyku Evrelerinin Uyku Döngüsü Boyunca Dağılımı

Sağlıklı bir insan uyku döngüsü genellikle 80 dk'lık NREM uykusuyla başlar ve bu süreyi 10 dk'lık REM uykusu takip eder. 90 dk'lık bu döngü gece boyunca dört veya altı defa olmak üzere kendisini yeniler. Bu döngüden herhangi biri herhangi bir sebeple bölünürse uyku kalitesi düşer. Şekil 1.6'da 8 saatlik bir uyku döngüsünün uyku evrelerine dağılımı gösterilmektedir.



Şekil 1.6. Uyku evrelerinin 8 saatlik uyku dönemindeki döngüsü [14]

2. MATERYAL VE YÖNTEM

Bu çalışmada alınan PSG kayıtları içerisinde 1 kanal EEG, 2 kanal EOG sinyali kullanılmış olup, bu verilerden öznelik çıkartımı yapılmıştır. Veriler filtrelendikten sonra dalgacık analizi yöntemiyle alt frekans bantlarına ayrılmış öznelikler frekans ve zaman domeninde hesaplanmıştır. Elde edilen veriler, her epok için bir matris içerisinde birleştirilmiş ve Weka programı dosyası olan Arff (Attribute Relationship File Format) dosya biçimine manuel olarak çevrilmiştir. Weka programı aracılığıyla makine öğrenmesi teknikleri kullanılarak özellik azaltımı gerçekleştirilmiş ve farklı sınıflandırıcılara öznelikler uygulanmıştır. Toplam 14542 adet epokun %20'si eğitim, %80'i test amacıyla kullanılmış ve en yüksek başarı oranı hesaplanmaya çalışılmıştır.

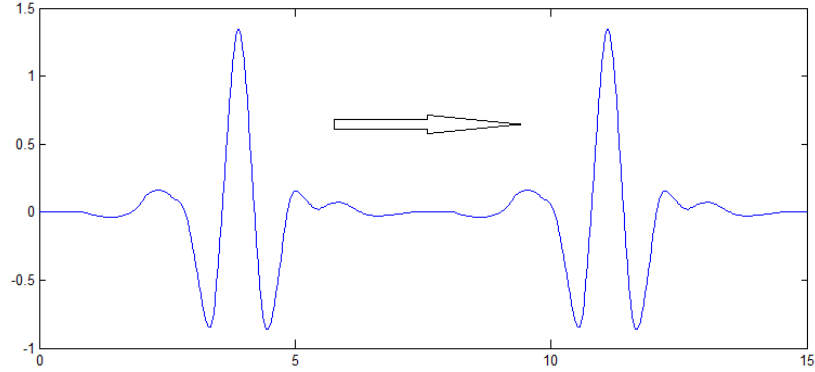
2.1. Özellik Çıkartımı

Tez çalışmasında EEG verilerini alt frekans bantlarına ayırmak için, ayrık dalgacık dönüşümü kullanılmıştır. Elde edilen alt bantlardan özellik çıkarımları yapılmıştır.

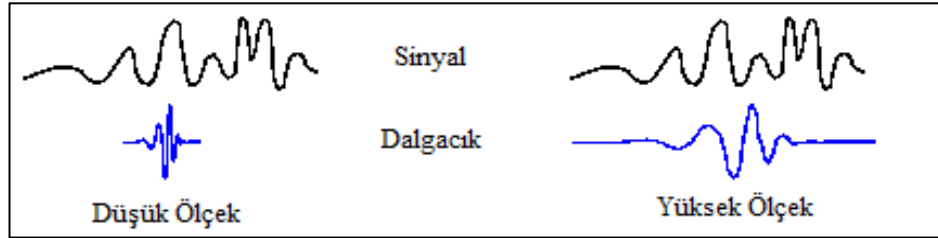
2.1.1. Dalgacık dönüşümü

Bir dalgacık, ortalama değeri sıfır olan ve zamanla sınırlı bir dalga şeklindedir. Dalgacıklar, zamanda uzatılarak veya kısaltılarak ölçeklendirilebilirler ve temel şekillerini değiştirmeden herhangi bir ölçekte farklı zaman konumlarına hareket ettirilebilirler. Aynı temel dalgacık şekline sahip bütün ölçeklendirilmiş dalgacıklar bir aile oluştururlar. Bir dalgacık ailesinde her hangi bir dalgacık için sonsuz sayıda ölçek ve ötelenme elde edebildiğimiz için sonsuz sayıda dalgacık bulunur. Şekil 2.1'de Sym 8 (Symlet) dalgacığının zamanda ötelenmesi gösterilmektedir. Şekil 2.2'de yüksek ve düşük ölçekli dalgacıklar gösterilmiştir [15]. Düşük ölçek sıkıştırılmış dalgacık, hızlı değişimleri daha iyi tespit eder. Yüksek frekanslı bileşenleri tespit etmede daha uygundur. Yüksek ölçek genişletilmiş dalgacık, yavaş değişimleri daha iyi tespit eder. Düşük frekanslı bileşenleri tespit etmede daha uygundur.

Bu durum dalgacık analizlerine, EEG, EOG gibi çok küçük ölçeklendirilmiş sinyalleri tespit etmek için gerekli olan çözünürlüğü sağlar [16].



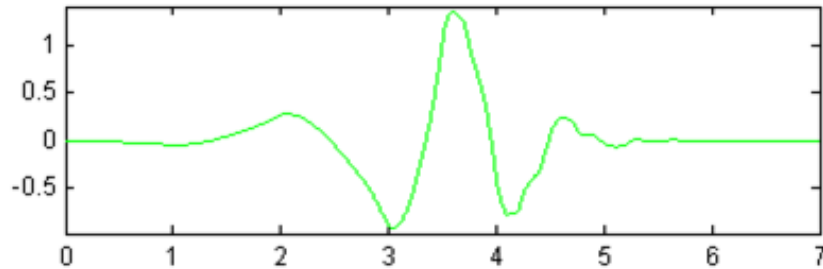
Şekil 2.1. Dalgacığın zamanda ötelenmesi



Şekil 2.2. Dalgacığın ölçeklenmesi

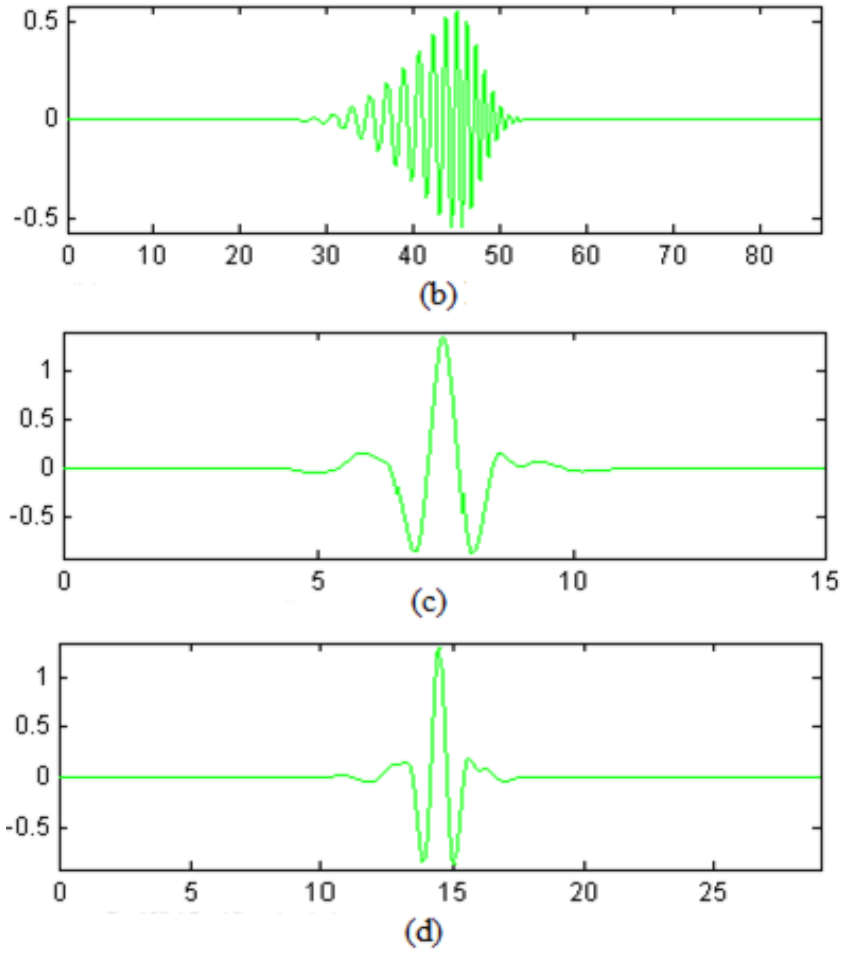
Dalgacıkların pek çok farklı biçime sahip olabilecekleri göz önüne alındığında, aradığımız sinyale benzeyen bir şekle sahip olan dalgacığı kullanmak mümkündür. Uygun seçilen dalgacıklar EEG sinyalinde bulunan özel dalga şeklinin belirli bir bileşeninin geçici ve spektral özelliklerini yakın olarak modelleyebilir. Böylelikle aradığımız sinyal değişimini yakalayabilecek duruma gelebiliriz [16].

Bazı dalgacık çeşitleri şunlardır :



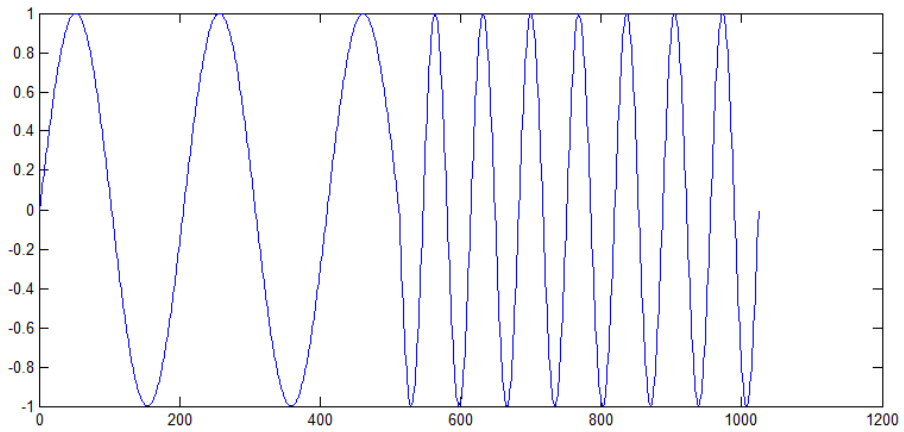
(a)

Şekil 2.3. Dalgacık çeşitleri. (a) Db4 (Daubechies Dalgacığı)
(b) Db 44 (c) Sym 8 (d) Coif 5 (Coiflets)

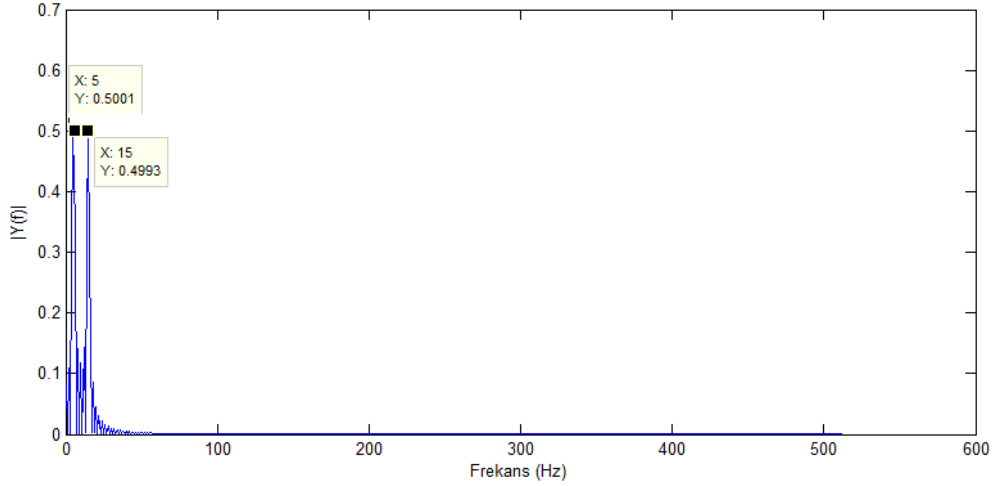


Şekil 2.3.(Devam) Dalgacık çeşitleri. (a) Db4(Daubechies Dalgacığı) (b) Db 44 (c) Sym 8 (d) Coif 5 (Coiflets)

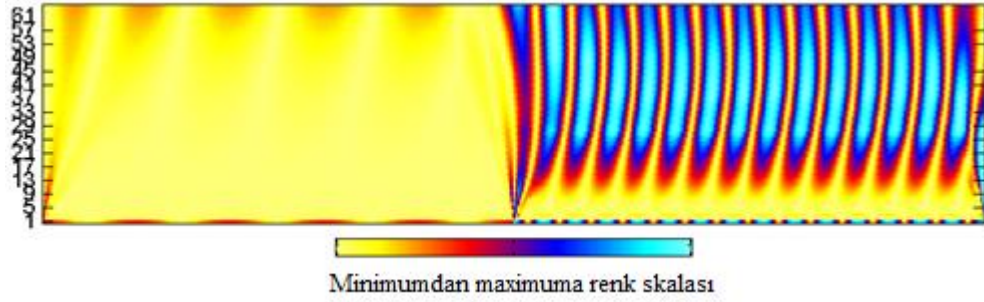
Fourier dönüşümünün durağan olmayan sinyallerde zaman frekans çözünürlüğü gibi dezavantajları gidermek için dalgacık dönüşümü geliştirilmiştir [11]. Canlılardan elde edilen sinyallerin hemen hepsi durağan olmayan sinyaller olduğu için frekans bilgisine ihtiyaç duyulmaktadır [17].



Şekil 2.4. Zamana göre frekansı değişen sinüs dalgası



Şekil 2.5. Zamana göre frekansı değişen sinüs dalgasının fourier dönüşümü



Şekil 2.6. Zamana göre frekansı değişen sinüs dalgasının dalgacık spektrumu

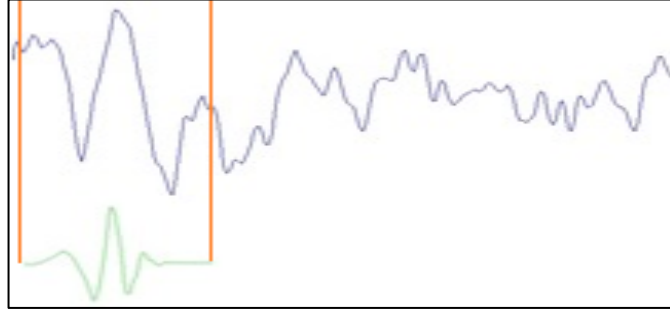
Şekil 2.4'de zamana göre frekansı değişen bir sinüs görülmektedir. Şekil 2.5'de ise bu fonksiyonun fourier dönüşümü verilmiştir. Bu dönüşüm sonucunda 5 Hz ve 15 Hz'lik 2 tane tepe ile sinyalin gösterildiği görülmektedir. Şekil 2.5'de zaman bilgisi yoktur. Dalgacık dönüşümünde ise zaman-ölçek ortamında sinüs dalgasının incelenebildiği Şekil 2.6'da gösterilmektedir.

Dalgacık dönüşümü temel olarak 3 ana başlık altında toplanabilir.

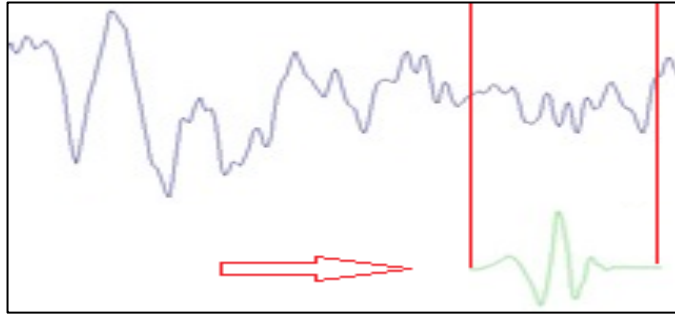
2.1.1.1. Sürekli dalgacık dönüşümü

Sürekli dalgacık fonksiyonunda analiz edilen fonksiyon bir dalgacıktır. Elde edilen dalgacık, zaman ekseninde ölçeklenip işlemek istediğimiz sinyal üzerinde kaydırılarak, korelasyon değerine bakılmasıyla aranan dalgacığın sinyal üzerinde bulunmasını sağlar. Çeşitli ölçeklerde ve ötelenmiş şekillerde dalgacık sinyali karşılaştırarak, iki değişkenli bir fonksiyon elde edilir. Ölçek, düzenlilik hakkında

bilgi verirken ötelenme ise dalgacığın oluşum anı hakkında bilgi verir. Ölçekleme işlemi, bir dalgacığı yatay ekseninde belli bir oranda gezdirmek demektir. Ölçek büyüdükçe dalgacık da aynı oranda genişletilir [11,18].



(a)



(b)

Şekil 2.7. Sürekli dalgacık dönüşümü. (a) Sürekli dalgacık dönüşümü zamanda kaydırma öncesi (b) Sürekli dalgacık dönüşümü zamanda kaydırma sonrası

Şekil 2.7(a) örneğindeki korelasyon katsayısının Şekil 2.7(b)'deki korelasyon katsayısından daha büyük olması beklenmektedir. Çünkü dalgacık formu ana sinyale daha çok benzemektedir.

Sürekli dalgacık yönteminin temel algoritması aşağıdaki adımlardan oluşmaktadır.

1. Bir dalgacık seçilir ve karşılaştırma yapmak için orijinal sinyalin başlangıcından başlanır.
2. C katsayısı hesaplanır. C katsayısı orijinal sinyal ile seçilen dalgacık arasındaki benzerlik katsayısıdır. C değerinin büyük olması dalgacık ile orijinal sinyalin birbirine benzer olması anlamına gelmektedir.

3. Zamanda dalgacık bütün sinyali kapsayana kadar kaydırılır ve 2 nolu adım her seferinde tekrar edilir.

4. Dalgacık ölçeklenerek 2 ve 3 nolu adımlar tekrar edilir [19].

Sürekli dalgacığın matematiksel ifadesi [19],

$$SDD(\tau,s)=\frac{1}{\sqrt{|s|}}\int_{-\infty}^{+\infty}f(t)\varphi\left(\frac{t-\tau}{s}\right)dt \quad (2.1)$$

biçimindedir. Buradaki s ölçek ve τ (to) ötelenmeyi, $f(t)$ dönüşümü yapılacak sinyali, φ ise dalgacığı ifade etmektedir. SDD'nin sonucunda ölçek ve ötelenme fonksiyonu olan birçok dalgacık katsayıları olan korelasyon değeri C elde edilir. Benzerlik, benzer frekans bileşenlerinin algılanması anlamına gelmektedir. Her olası ölçek için dalgacık katsayılarının hesabı gereksiz birçok bilgi üretilmesine ve fazladan işlem yapılmasına neden olur. Bu nedenle ayrık dalgacık dönüşümü kullanılması daha uygun olacaktır.

2.1.1.2. Ayrık dalgacık dönüşümü

Ayrık dalgacık dönüşümüne olan gereksinim sürekli dalgacık dönüşümü bir işarete uygulandığında işaretle ilgili olarak elde edilen dalgacık katsayılarının, her olası ölçek için gereksiz birçok bilgi üretmesine neden olur. Bu durum ise bilgisayar belleğinde çok yer kaplar hem de gereğinden fazla zaman alır. Ayrık dalgacık dönüşümü bu dezavantajları azaltmak için kullanılır. Matematiksel olarak sürekli dalgacık dönüşümü ile ayrık dalgacık dönüşümü benzerdir [20]. Ayrık dalgacığın matematiksel ifadesi [20],

$$ADD(s,\tau)=2^{-s/2}\int f(t)\varphi\left(2^{-s}t-\tau\right)dt \quad (2.2)$$

Burada $f(t)$ işlenecek sinyali, φ parametresi dalgacığı, s parametresi ölçeklemeyi, τ parametresi ise ötelemeyi göstermektedir. Düşük frekanslı bileşenleri incelemek için büyük öteleme değeri, yüksek frekanslı bileşenleri incelemek için ise küçük öteleme değeri kullanılır.

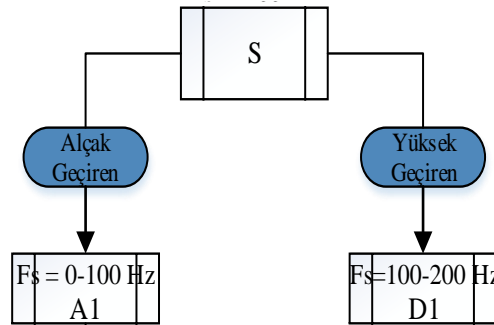
Ayrık dalgacık dönüşümünde orijinal sinyali farklı frekans aralıklarına ayırmak için filtreler kullanılır. En yüksek frekans bileşeni f olan orijinal sinyal, birbirini

tamamlayan yüksek geçiren ve alçak geçiren filtrelerden geçirilir. Bu işlem istenilen frekans aralığına ulaşana kadar çok kez tekrarlanabilir. Yüksek geçiren filtre çıkışı ayrıntılı (D), alçak geçiren filtre çıkışı Yaklaşım(A) katsayılarını verir [11]. ADD(Ayrık dalgacık dönüşümü)'nde ölçekleme ve öteleme değerleri tespit edilen ölçek aralıklarında 2'nin katlarını temel alacak şekilde seçilir [21].

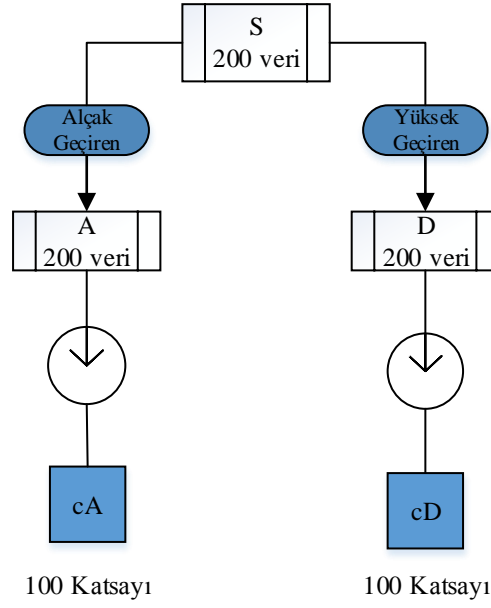
Filtrelemeden sonra alçak geçiren filtrede en yüksek frekans bileşeni f yerine $f/2$ olur ve Nyquist kriterlerine göre örneklerin yarısı elenir. Nyquist örnekleme frekansı, sinyalin tekrar oluşturulmasında bilgi kaybına uğramamak için orijinal sinyalin en büyük frekansından iki kat büyük seçilen sınıra denir. Bu frekansın altında yapılacak örneklemler, sinyalin yeniden oluşumunda "aliasing" denilen bilgi kayıplarına sebep olur.

Örnek ile açıklayacak olursak;

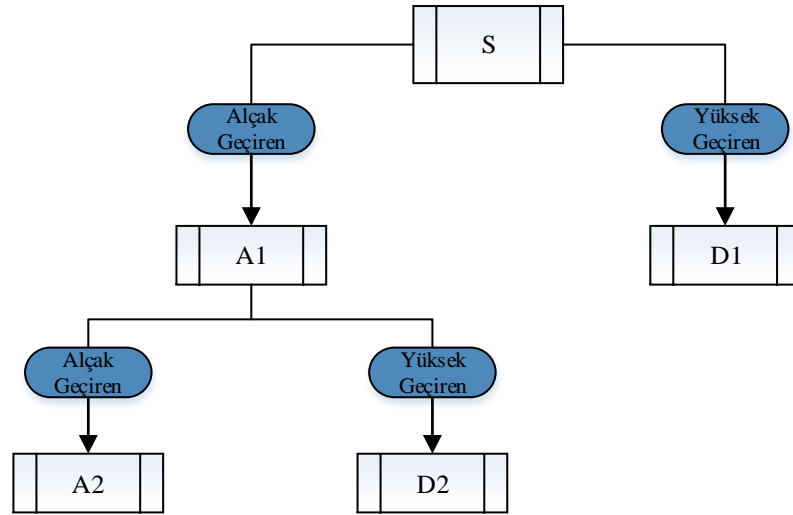
S sinyalimiz 200 adet veri bulunan bir zaman serisi ise S sinyalini birbirini tamamlayan alçak ve yüksek geçiren filtreden geçirdiğimizde toplamda 400 adet veri ortaya çıkar. Şekil 2.8'de S sinyalinin alçak ve yüksek geçiren filtreden geçirildikten sonra elde edilen alt frekans bantları gösterilmektedir. Orijinal sinyalin kendisiyle aynı sayıda veri yerine 2 katı veri ile temsil edilmesi istenilen bir durum değildir. Bu durumda orijinal veri miktarına dönebilmemiz için aşağı örnekleme işlemi yapılır. Bu işlem Şekil 2.9'da gösterilmiştir. Bu işlem sinyalin zaman çözünürlüğünü azaltır. Çünkü örnek sayısı yarı yarıya düşmüştür. Bunun diğer bir anlamı zaman ekseninde yarı yarıya bilgi kaybı meydana gelir. Frekans çözünürlüğü ise artar. Bunun sebebi ise bir önceki duruma göre frekans çeşitliliği yarı yarıya azaldığından kalan frekanslar daha iyi tespit edilebilir [22].



Şekil 2.8. Bir zaman serisinin aşağı örnekleme algoritması



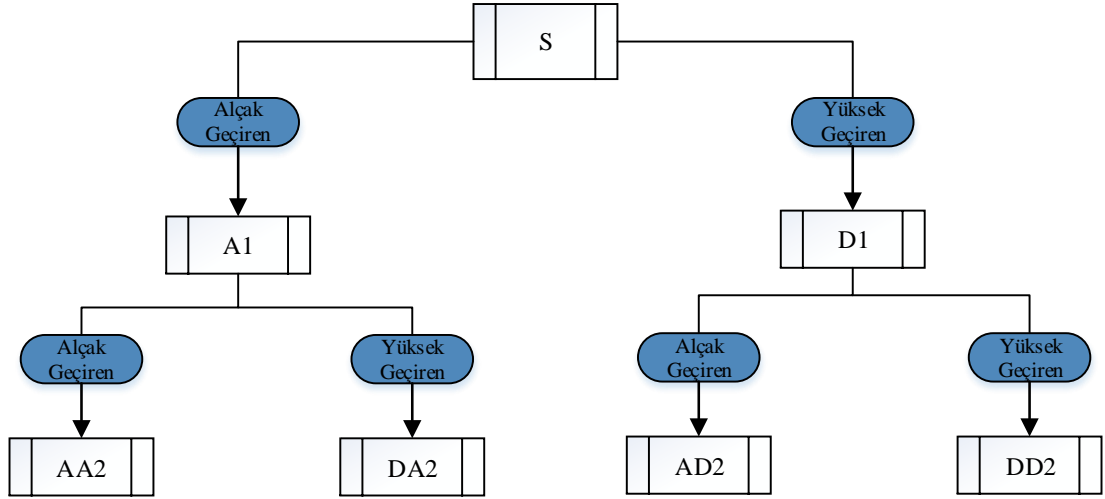
Şekil 2.9. Gerçek zamanlı serinin aşağı örneklenmesi



Şekil 2.10. 200 örnekli bir sinyalin 2.seviye ayrışım basamakları

2.1.1.3. Dalgacık paket dönüşümü

Ayrık dalgacık dönüşümünde sadece yaklaşım katsayıları alt bantlara ayrıştırılırken, dalgacık paket dönüşümünde hem yaklaşım hem de detay katsayıları alt bantlara ayrıştırılarak daha detaylı işaret işleme olanağı sağlanmaktadır [11]. Sinyalin 2. seviyeden bileşenlerine ayrıştırılması Şekil 2.11'de gösterilmiştir.



Şekil 2.11. Paket dalgacık dönüşümü 2. seviye ayrışım basamakları

2.1.1.4. Dalgacık dönüşümünün kullanım alanları

Dalgacık dönüşümünün sayısal sinyal işleme teknikleri içerisinde iki temel kullanım alanı vardır. Bunlardan biri veri sıkıştırma, ikincisi gürültü gidermedir. Bunların dışında da birçok kullanım alanı vardır. Örnek olarak matematik alanında, görüntü işleme, kuantum mekaniği, teorik fizik, askeri elektronik harp ve silah zeka, bilgisayar sınıflandırılması ve tanımlanması; yapay müzik ve dil sentezi, tıbbi görüntüleme ve tanı vb.. Veri sıkıştırmanın elverişli olmasının sebebi dalgacık dönüşümünün yapısından kaynaklanmaktadır. Çünkü kesikli dalgacık dönüşümü ayrıştırma işlemi sırasında her seviyede örnek sayısını azaltılıp, fazlalık bilgi elediği için, orijinal sinyale göre tekrar sentezlenmiş hali oldukça az yer kaplar. Bu yöntem kullanılan değişik kodlama teknikleri ile gerçekleştirilebilir [18-22].

2.1.2. Enerji

Enerjisi sıfırdan farklı ve sonlu olan sinyallere enerji sinyali denir. Bir gerçel ya da karmaşık sürekli $f(t)$ fonksiyonu [16],

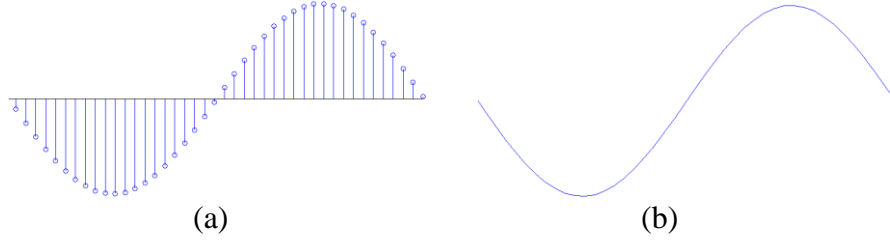
$$\int_{-\infty}^{+\infty} |f(t)|^2 dt < \infty \quad (2.3)$$

Eşitlik (2.3)'deki eşitlik sağlanıyorsa bu $f(t)$ fonksiyonuna "enerji işareti" denir. Verilen integral ifadesi işarettaki enerjiyi vermektedir. Farklı bir söylemle, bir

işaretin enerjisi $(-\infty, \infty)$ aralığında sonlu bir değer ise bu işarete enerji işareti denir. Benzer şekilde, ayrık $f[n]$ işareti [16],

$$\sum_{-\infty}^{+\infty} |f[n]|^2 < \infty \quad (2.4)$$

şartını sağlıyorsa bu ayrık $f[n]$ işareti enerji işareti olarak adlandırılır. Aşağıdaki Şekil 2.12’de tipik enerji işaretleri verilmiştir.



Şekil 2.12. Ayrık ve sürekli enerji işaretleri. (a) Ayrık İşaret (b) Sürekli işaret

Dalgacık dönüşümünden elde edilen her bir alt frekans bandı için enerji hesabı yapılmıştır. Alt bantlar için elde edilen enerji değerlerinin birbirlerine oranları da öznitelik olarak kullanılmıştır.

2.1.3. Maksimum tepe değeri

İşlenmeye hazır hale getirilen EEG sinyali içerisinde en yüksek genlik değerinin hesaplanmasıdır. Dalgacık yöntemiyle elde edilmiş olan alt frekans bantları için her alt bandın içerisindeki maksimum genlik değeri hesaplanmıştır.

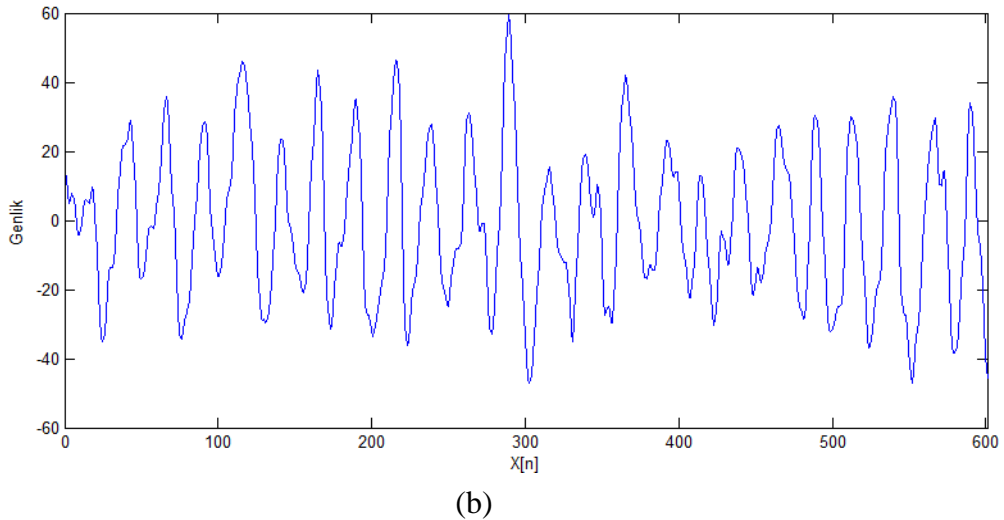
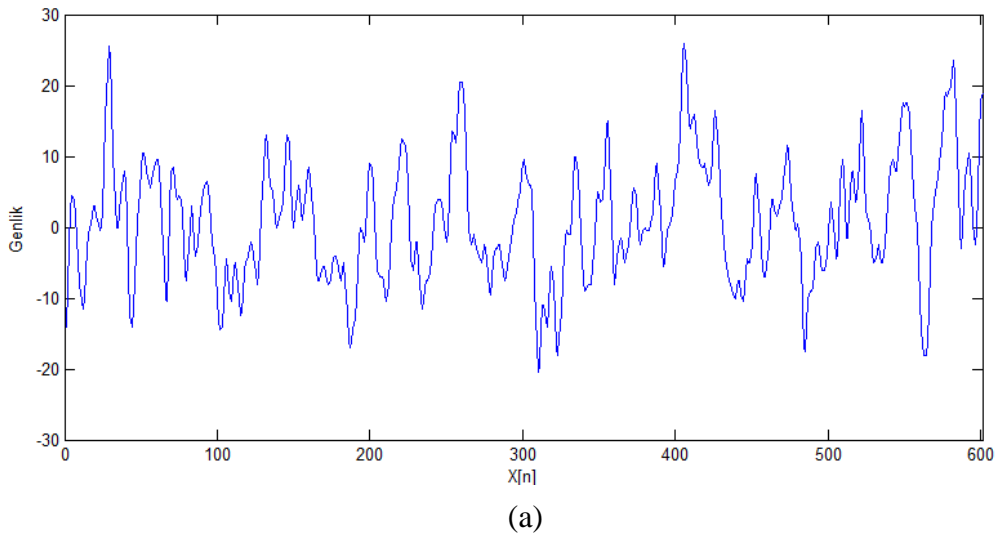
2.1.4. Minimum tepe değeri

İşlenmeye hazır hale getirilen EEG sinyali içerisinde en düşük genlik değerine sahip olan değer tespit edilmesidir. Dalgacık yöntemiyle elde edilmiş olan alt bantlar için her alt frekans bandının içerisindeki minimum genlik değeri hesaplanmıştır.

2.1.5. Standart sapma

Standart sapma varyansın kareköküdür. Dalgacık dönüşümü ile elde edilmiş olan alt bantların her biri için ayrı ayrı standart sapma değerleri hesaplanmıştır. Geniş salınımlı bir EEG işareti yüksek varyans değerleri verirken sabit bir EEG işareti düşük varyans değerleri verir. Standart sapmanın matematiksel ifadesi [18],

$$\sigma = \sqrt{\frac{1}{N} \sum_{n=1}^N (X_n - \mu_x)^2} \quad (2.5)$$



Şekil 2.13. EEG kaydı. (a) 3 saniyelik düşük genlikli EEG kaydı (b) 3 saniyelik yüksek genlikli EEG kaydı

Şekil 2.13'de aynı hastadan farklı zamanlarda kaydedilmiş 3'er sn'lik EEG kaydı gösterilmektedir. Şekil 2.13(a)'da daha yavaş dalgalar ve düşük genlikler sinyalin

geneline hakim iken Şekil 2.13(b)'de daha hızlı dalgalar ve yüksek genliklerin sinyalin geneline hakim olduğu görülmektedir. Şekil 2.13(a)'daki sinyalin standart sapması hesaplandığında 8,52 , Şekil 2.13 (b)'deki sinyalin standart sapması hesaplandığında 21,83 çıkmaktadır.

2.1.6. Ortalama değer

X, N adet örnekten oluşan bir öznitelik vektörü ise,

$$X[n]= x_1, x_2, \dots, x_N \text{ şeklindedir.} \quad (2.6)$$

Bu örneklerin ortalaması,

$$\mu_x = \frac{1}{N} \sum_{n=1}^N X_n \quad (2.7)$$

şeklinde bulunur [18].

Ortalama değeri ayrıca X'in beklenen değeri olarak ifade edilir ve Eşitlik (2.8)'teki şekilde gösterilir.

$$\mu_x = E\{X\} \quad (2.8)$$

2.1.7. Entropi

Entropi, bir rastsal değişkenin belirsizlik ölçüsüdür. Başka bir deyişle bir sistemin düzensizliğinin ölçüsüdür. Entropi ilk kez Shannon tarafından ortaya atılmıştır ve en çok bilinen entropi Shannon entropisidir [23]. Shannon'dan sonra Norm, Eşik, Logaritmik, Örnek ve Yaklaşım (Approximate) entropisi gibi entropi hesaplama metotları da geliştirilmiştir [24]. Bu çalışmada, her bir alt bant için Shannon entropi'si hesaplanmıştır.

Shannon entropisi,

$$H(X) = - \sum_x p(x) \ln p(x) \quad X \text{ sinyali ayrık ise} \quad (2.9)$$

$$H(X) = - \int f(x) \ln f(x) dx \quad X \text{ sinyali sürekli ise} \quad (2.10)$$

şeklinde tanımlanmaktadır [23,24].

Şekil 2.13'deki sinyallerin entropisi hesaplandığında;

Düzensizliğin daha az olduğu Şekil 2.13(a)'nın entropisi -1,8730 olarak hesaplanırken, düzensizliğin daha çok olduğu Şekil 2.13(b)'nin entropisi -60,91 olarak hesaplanmıştır. Bu durum düzgün bir olasılık dağılımına sahip olan verinin entropisi yüksek, düzensiz bir olasılık dağılımına sahip olan verinin entropisinin ise düşük olacağını göstermektedir.

2.1.8. Hjorth parametreleri

Hjorth tarafından, EEG işaretlerini zaman uzayında tanımlayabilmek için üç değişken hesaplanmıştır. Bunlar genel olarak Hjorth parametreleri olarak bilinir ve isimleri ise aktivite, hareketlilik ve karmaşıklık. Hjorth parametreleri hesaplanırken işaretin birinci ve ikinci türevlerinden faydalanılır [25].

İlk parametre olan aktivite değeri işaretin ortalama enerjisine eşittir. İkinci parametre olan hareketlilik değeri ortalama frekansın kestirimidir. Son parametre olan karmaşıklık ise frekans değişimidir [26]. Kullanacağımız sinyale n uzunluğunda $x(n)$ dersek,

$$\text{Aktivite(Activity)} = \text{var}(x(t)) \quad (2.11)$$

$$\text{Hareketlilik(Mobility)} = \sqrt{\frac{\text{Activity}\left(\frac{dx(t)}{dt}\right)}{\text{Activity}(x(t))}} \quad (2.12)$$

$$\text{Karmaşıklık} = \frac{\text{Mobility}\left(\frac{dx(t)}{dt}\right)}{\text{Mobility}(x(t))} \quad (2.13)$$

$\text{var}(x(t))$, x sinyalindeki varyans ve \bar{X} ortalama değer ise

$$\text{var}(x) = \frac{\sum_{n=1}^N (x(n) - \bar{X})^2}{N} \quad (2.14)$$

şeklinde bulunur [12,25].

2.2. Makine Öğrenmesi Teknikleri

Değişken seçimi, nitelik seçimi veya değişken bir alt küme seçimi olarak adlandırılan makine öğrenmesi, model çıkartımı kullanımı için uygun bir özellik alt-kümesinin seçilmesi işlemidir. Temel varsayım özellik seçim tekniği ile eldeki ayırt edici verilerin, gereksiz ya da alakasız özellikler içerebileceği düşüncesi temeline dayanmaktadır.

Girdiler kötü seçildiğinde kullanılan bu girdilerle, en iyi sınıflandırıcı kullanılsa bile istenen performans alınamaz. Bu nedenle, sınıflandırılacak veriyi en iyi temsil edecek şekilde girdi seçimi yapılmalıdır [22].

Özellik seçimi teknikleri ile özellik çıkarımı teknikleri birbirinden ayırt edilmelidir. Özellik çıkartma eldeki verilerden yeni bir özellik çıkartımıdır. Özellik seçimi ise özelliklerin bir alt kümesini oluşturur. Özellik seçimi teknikleri genellikle çok sayıda özellik (veri noktaları) olduğu durumlarda kullanılır. Özellik seçimi tekniklerinin modele getireceği yarar, daha kısa eğitim süresi ve modelin performansını arttırmasıdır. Bu çalışmada, özellik azaltımı EEG sinyallerinin frekans domeni özellikleri üzerinde istatistiksel ölçümler kullanılarak yapılmıştır.

2.2.1. Özellik azaltımı (Seçimi)

Özellik azaltımı genellikle iki kısımdan oluşur.

- Özellik alt küme üretici
- Değerlendirici

Özellik alt küme üretici durum uzayında arama yapar ve alt kümeleri oluşturur. Değerlendirici ise oluşturulmuş alt kümeleri değerlendirir ve ne kadar iyi olduklarını belirler. Bu iki kısım, değerlendirme kriterini en iyi karşılayan özellik alt kümesini bulmak için çalışır [27]. Bu çalışmada, kullanılan ranker arama motoru Relief, GainRatio gibi değerlendiricilerle birlikte kullanılarak her bir öznelik için bireysel değerlendirme sonucuna öznelikleri sıralayarak özneliklerin sınıflandırma başarımını tespit eder.

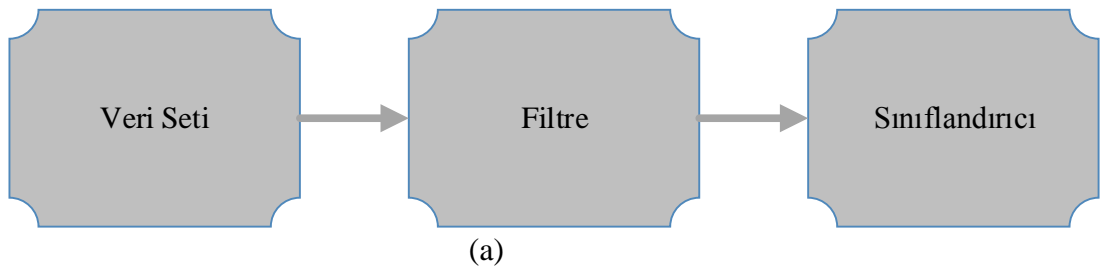
Özellik alt küme üretici, arama motoru gibidir. 3 kategoriye ayrılır;

- Ayrıntılı arama motoru (Exhaustive search engine)
- Sezgisel arama motoru (Heuristic search engine)
- Belirleyici olmayan arama motoru (Nondeterministic search engine)

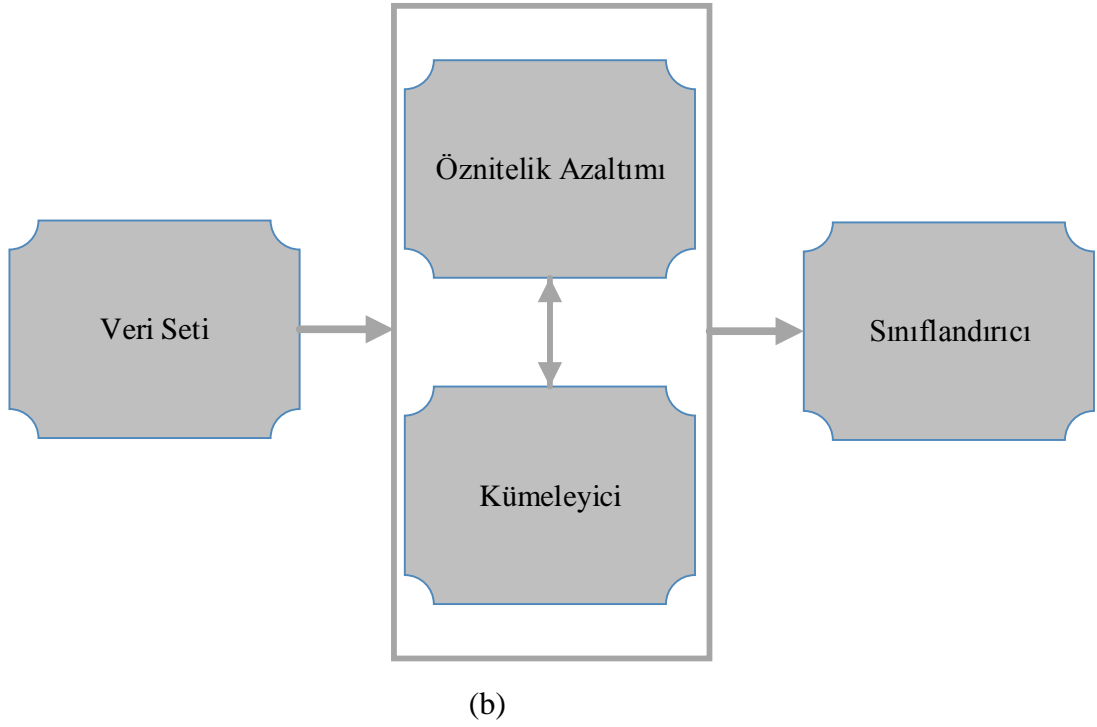
Bu arama motorları durum uzayında farklı arama stratejileri kullanarak arama yapar. Değerlendirici, bir öğrenme algoritmasına dahil olup olmamasına göre iki kategoriye ayrılır:

- Sarmal (Wrapper) metod
- Filtre metod

Sarmal metod değerlendirmek için sınıflandırıcıya özel öğrenme algoritması kullanarak ne kadar iyi bir alt kümeleri seçebileceğini değerlendirir. Aynı zamanda öznelikler birlikte değerlendirilerek bir alt küme oluşturulmaya çalışılır. Filtre metodu ise alt kümelerin ne kadar iyi olduğuna bakarak değerlendirme yapar. Şekil 2.14(a)'da görüldüğü her bir öznelik bilgisini bireysel değerlendirir. Kriterler puanlanarak sıraya koyulur ve bu puanlara göre kullanıcının seçtiği kadar ya da belirlenen eşik değeri üzerindeki öznelikler sınıflandırma için kullanılır. Filtrelerde değerlendirme için genellikle kullanılan ortak kriterler, korelasyon ve tutarlılıktır [28]. Öğrenme algoritmasına bağlı olmadan seçilen değerlendiricilere örnek olarak CfsSubsetEval, GainRatioAttributeEval, InfoGainAttributeEval verilebilir.



Şekil 2.14. Değerlendirici (a) Filtre metodu (b) Sarmal metod



Şekil 2.14.(Devam) Değerlendirici (a) Filte metodu (b) Sarmal metot

Sınıflandırma algoritmalarındaki genel amaç sınıflandırma doğruluğunu arttırmaktır. Sarmal metot filtre metoduna göre genellikle sınıflandırma doğruluğunu daha yüksek oranlara taşıyabiliyor. Bu durum filtre metodunun öğrenme algoritmasını temel almadan değerlendirme yapmasından kaynaklanmaktadır. Şekil 2.14(b)'de görüldüğü gibi özellik azaltımı kümelenecek sınıflandırıcıya göre yapılır. Sarmal metot ise filtre metoduna göre genellikle daha yavaştır. Hız açısından değerlendirecek olursak filtre metodunun daha başarılı olduğu görülmüş ancak sınıflandırma açısından değerlendirildiğinde sarmal metodun genelde daha başarılı olduğu görülmüştür [27,29]. Bu çalışmada, filtre metodu ve sarmal metot ayrı ayrı değerlendirilmiştir.

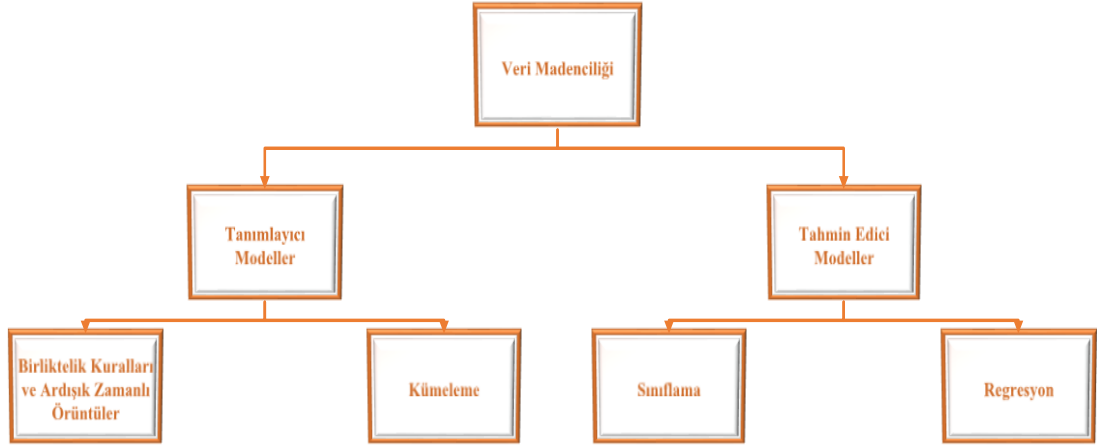
Tablo 2.1. Örnek öznitelik azaltımı tablosu

ELDE EDİLEN ÖZİNİTELİKLER	SINIFLANDIRICI	BAŞARI YÜZDESİ
X,Y,Z	A	%87
X,Y	A	%87
X,Z	A	%75
Y,Z	A	%78
X	A	%83
Y	A	%84
Z	A	%60

Tablo 2.1’de örnek bir öznitelik azaltım tablosu oluşturulmuştur. A sınıflandırıcısı kullanılarak X,Y ve Z özniteliklerinden elde edilen başarı oranıyla; X,Y özniteliklerinden elde edilen başarı oranı aynı olduğuna göre X,Y öznitelik çiftinin kullanılmasıyla elde edilen sonucun daha başarılı olduğu söylenebilir. Yani, özellik azaltarak daha az veriden daha yüksek başarı elde edilmesi hedeflenmiş oluyor.

2.3. Veri Madenciliği

Büyük veri yığınları içerisinde gelecek ile ilgili tahminde bulunabilmemizi sağlayabilecek bağıntıların bilgisayar programı aracılığıyla aranmasına veri madenciliği denir. Veri madenciliğini şekilde gösterildiği gibi alt başlıklara ayrılmaktadır.



Şekil 2.15. Veri Madenciliği

Veri madenciliği fonksiyonlarına göre tahmin edici (Predictive) ve tanımlayıcı (Descriptive) olmak üzere iki alt başlığa ayrılmaktadır [30]. Tahmin edici modellerde, öncelikle sonuçları bilinen verilerden bir model oluşturulur. Elde edilen bu model sonuçları bilinmeyen verilerin, sonuçlarının tahmin edilmesi amacıyla kullanılır. Örneğin bir bölgedeki elektrik tüketiminin, ileriki zaman aralıkları için tahmin edilip, bu tahminlere göre elektrik üretim firmalarından talepte bulunulması gerekmektedir. Çünkü elektrik enerjisi depolanabilir bir enerji değildir. Gelecekte ne kadarlık elektrik tüketilebileceğini hesaplayabilmek için geçmişteki benzer koşullarda (Hava şartları, nüfus vs..) ne kadarlık tüketim yapıldığı bilgisini kullanarak, tahmin edici veri madencilik modelleriyle yaklaşık olarak hesaplanabilir.

Tanımlayıcı modellerde ise karar vermede etkili olan mevcut verilerdeki ilişkilerin tanımlanması ve ne anlama geldiğini anlamayı amaçlar.

2.3.1. Birliktelik kuralı

Birliktelik kuralları, bir veri setinde yer alan özniteliğin değeri ile farklı bir veri setinde yer alan özniteliklerin değerleri arasında bir ilişki kurup gruplandırılması ile bulunur. Bu kurallar ilk olarak Agrawal tarafından 1993 yılında ortaya atılmıştır.

Birliktelik kuralı şu şekilde tanımlanabilir;

$$A1 , A2 , A3 ,An = B1 , B2 , B3, Bn \quad (2.15)$$

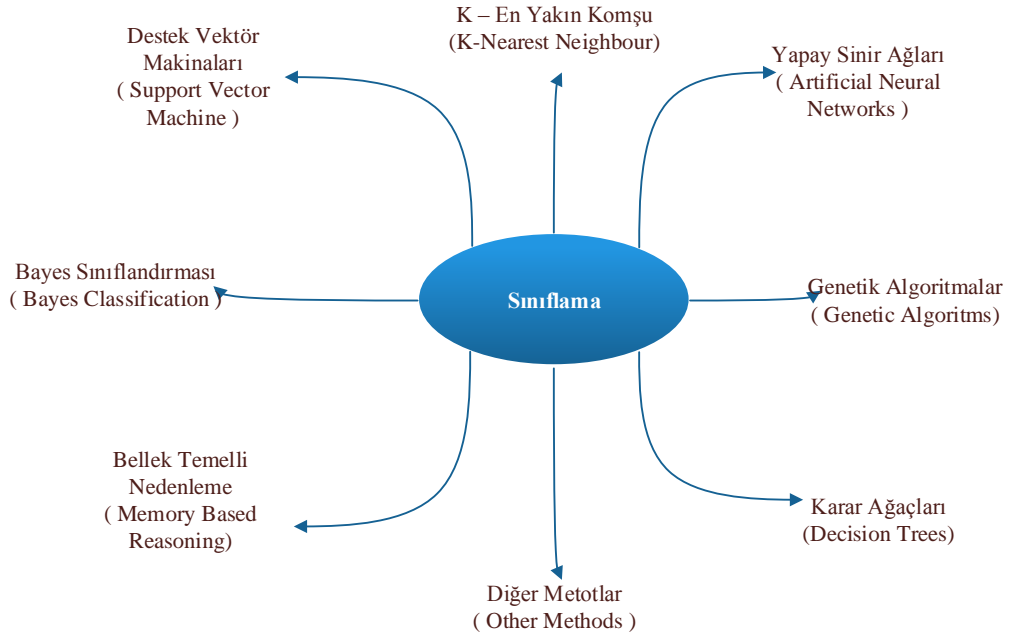
Bu ifadede yer alan A ve B yapılan iştir. Kural, genel olarak A işleri meydana geldiğinde, genel olarak B işlerinin de aynı olay içerisinde yer aldığını belirtir. Birliktelik kurallarının analiz süreci market sepeti analizi olarak da adlandırılır. En yaygın birliktelik kuralı algoritmaları Apriori sayılabilir [31,32].

2.3.2. Kümeleme modelleri (Clustering)

Kümeleme modelindeki amaç kümedeki veriler arasında özellikleri birbirinden çok farklı olan verileri bulması ve farklı kümelere ayrıştırmasıdır. Böylece, veriler örneklenen veri özelliklerini iyi yansıtan etkili temsil gücüne sahip olmaktadır [31,33].

2.3.3. Sınıflandırma (Classification) ve regresyon (Regression) modelleri

Tahmin için kullanılan ve veri madenciliği teknikleri arasında en yaygın kullanıma sahip olan modeldir. Sınıflandırmadaki amaç benzer özellikteki verilerin önceden belirlenmiş gruplara atanması ve sadeleşmesidir. Yani önceden belirlenmiş bir sınıf ve bu sınıfın eğitilerek bilinmeyen bir verinin bu sınıflardan birine dahil edilmesidir. Regresyon modelinde ise verinin diğer verilerle arasında bir ilişki olduğu kabul edilir. Sınıflama ve regresyon modellerinde karar ağaçları, yapay sinir ağları, genetik algoritmalar, K-en yakın komşu ve naive-bayes gibi teknikler kullanılmaktadır. Şekil 2.16'da sınıflama modelinde kullanılan yöntemler gösterilmiştir.



Şekil 2.16. Sınıflandırma modelinde kullanılan başlıca yöntemler

2.3.3.1. KNN komşuluğu (KNN neigbohours)

KNN sınıflandırıcı tipik sinyalleri sınıflandırmak için birçok araştırmacı tarafından kullanılan basit ve sezgisel bir yöntemdir. Bu sınıflandırıcı test verileri ile eğitim verileri arasında bir karşılaştırmaya karar verir. İki karar kuralları vardır, benzerlik ve çoğunluk [34].

Sınıflandırma için kullanılan bu algoritmada, sınıflandırma sonucunun tahmini için kullanılacak olan yeni veriye, daha önceden öznitelikleri elde edilmiş ve bu öznitelikler ile sınıflandırma sonucu bilinen verilerden, K tanesinin yakınlığına bakılmasıdır.

Örneğin $K = 5$ için yeni bir veri sınıflandırılmak istendiğinde, eğitim verisinde sınıflandırılmış olan verilerden en yakın 5 tanesi seçilir. Bu veriler arasında hangi sınıf baskın ise, yeni veri de baskın sınıfa dahil edilir. Eşit kalma durumunu önlemek için ise tek sayılarda (1,3,5 gibi) komşuluk kullanılır. Mesafe hesabında genelde öklit mesafesi kullanılabilir [35].

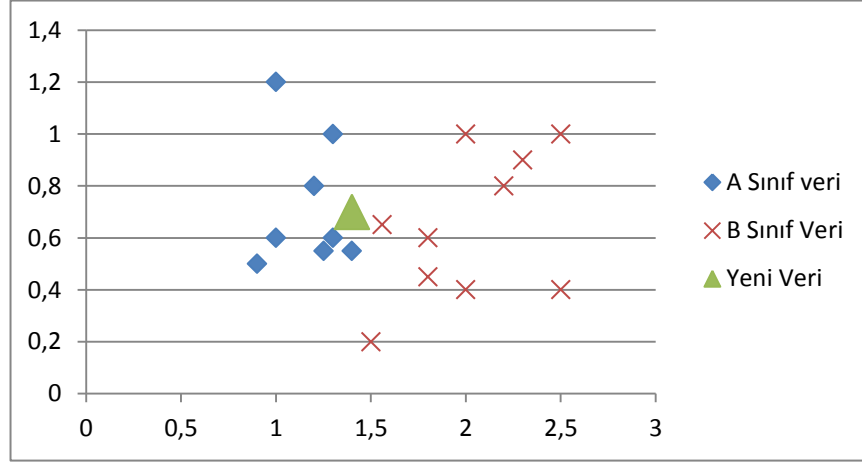
Öklid Mesafe hesabı,

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_i - y_i)^2} \quad (2.16)$$

i = Öznitelik sayısı

x = Test kümesine ait öznitelikler kümesi

y = Eğitim kümesine ait öznitelikler kümesi



Şekil 2.17. KNN sınıflandırıcısı

Algoritma, yeni bir örnek geldiğinde, onun en yakın K komşusuna bakarak örneğin sınıfına karar verir.

Şekilde yeşil renk ile gösterilen yeni bir veri, veri havuzuna geldiğinde KNN algoritması aşağıdaki aşamaları gerçekleştirecektir.

1. K değerine bak
2. Tüm bilinen örneklerin yeni örneğe olan uzaklığını hesapla
3. Minimum uzaklığa göre sırala
4. Ait oldukları sınıf değerlerini bul
5. Değeri baskın olan sınıfı seç

2.3.3.2. Destek vektör makinası (Support vector machine)

SVM 1960'lı yıllarda V. Vapnik tarafından teorik olarak ortaya atılmış ve zaman içerisinde sınıflandırma ve regresyon problemleri çözümü için önerilmiştir [36]. Bu metot, başlarda iki sınıflı problemlerin çözümünde doğrusal bir sınıflayıcı olarak düşünülmüş, daha sonra doğrusal olarak ayrılamayan ya da çok sınıflı sınıflama problemlerinin çözümünde de genelleştirilerek kullanılmaya başlanmıştır [37]. Destek vektör makineleri herhangi bir sınıflandırma problemini, bir karesel programlama problemine dönüştürerek yerel minimumlara takılmadan çözer [38].

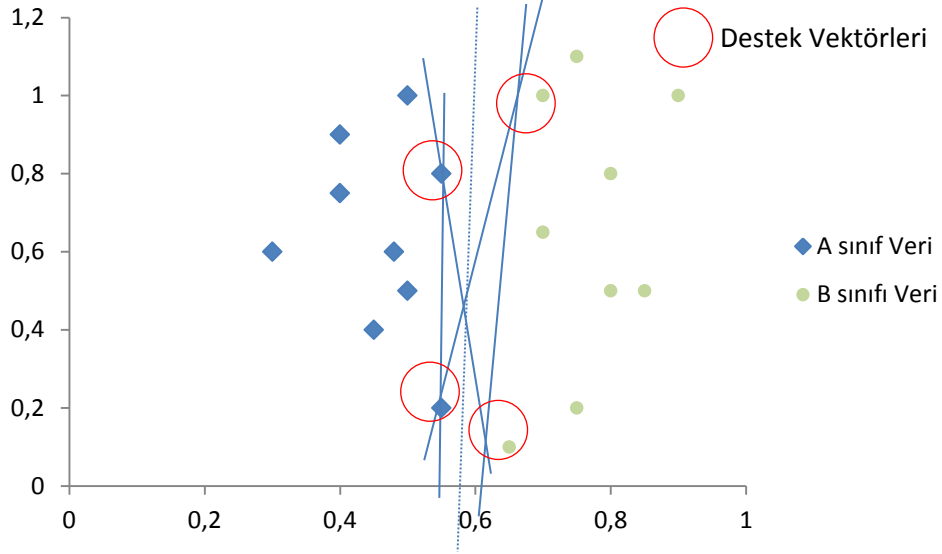
SVM’de amaç, iki sınıf arasındaki uzaklığın maksimum olduğu durumun bulunmasıdır. Bu verileri iki sınıfa ayıracak çok fazla sayıda ayırıcı düzlem bulunmaktadır. Optimum ayırıcı düzlem, her iki sınıfın en uç verileri arasındaki mesafenin (iki sınıfın destek vektörleri arasındaki mesafe) maksimum olduğu durumu sağlayan ayırıcı düzlemdir [36]. SVM’yi oluşturan başlıca bileşenler,

- İstatiksel öğrenme
- Optimizasyon algoritması
- Çekirdek fonksiyonlarıdır.

SVM’nin günümüzde sıklıkla kullanılıyor olmasının temel nedenlerinden bir tanesi uygun kernel fonksiyonu seçimi ile karmaşık düzgün olmayan verileri modelleme başarısıdır. Bu durum da yüksek boyutlu bir özellik uzayında doğrusal olmayan ilişkilerin, doğrusal bir çizgi olarak ifade edilmesi anlamına gelmektedir. Diğer bir deyişle, doğrusal olarak verilerin ayırt edilemediği durumlarda verileri daha büyük boyutlu uzaylara taşımak için bir fonksiyon kullanarak boyut dönüşümleri yapılabilir. Bu fonksiyonlara çekirdek fonksiyon denir. Bu amaçla çekirdek fonksiyonlardan yararlanır [17]. Destek Vektör makinaları doğrusal ve doğrusal olmayan olarak ikiye ayrılır.

Doğrusal destek vektör makinaları, birbirinden doğrusal olarak ayrılabilen veriler için kullanılır. Bir hiper düzlem, üzerindeki noktalar cinsinden aşağıdaki denklem ile elde edilebilir. Burada, x öznitelik, b sabit, w ise ağırlık vektörü olmak üzere,

$$w \cdot x + b = 0 \quad (2.17)$$



Şekil 2.18. Destek vektör makinesi

Şekil 2.18 üzerinde veriler farklı ve çok sayıda doğru ile ayrılabilir. Çok boyutlu uzayda bu doğruların yerini hiper düzlemler almaktadır. Bu iki hiper düzlemin ortasını oluşturan kesikli çizgilerle gösterilen düzlem ise iki sınıftaki verileri birbirinden ayıran doğrusal hiper düzlemdir. Bu düzleme en uygun ayırma hiper düzlemi denir.

Doğrusal olmayan SVM’de, veriler düz bir hiper düzlem ile birbirinden ayrılamazlar. Bu durum için çekirdek fonksiyonlar kullanılır ve en uygun ayırma düzlemine bu çekirdek fonksiyonları ile ulaşılır.

Destek vektör makinasının başarımı kernel fonksiyonuna ve çekirdek fonksiyona bağlı değişken parametrelere bağlıdır. SVM ile gerçekleştirilecek bir sınıflandırma işlemi için kullanılacak çekirdek fonksiyonu ve bu fonksiyona ait en uygun parametrelerin belirlenmesi esastır.

5 tane çekirdek fonksiyonu vardır. Bunlar [39] :

- ❖ Doğrusal(Linear),
- ❖ Polinomial(Polynomial),
- ❖ Karesel (Quadratic),
- ❖ RBF(Radial basis function) çekirdek fonksiyonu
- ❖ PUK(Pearson VII Universal) çekirdek fonksiyonu

Doğrusal (linear) çekirdek fonksiyonu, SVM’de çıkış değeri üretiminde en etkili olan çekirdek fonksiyonudur. Eşitlik (2.18)’de gösterildiği gibi x ve y ’nin çarpımı ile hesaplanan fonksiyondur [39].

$$k(x_i, y_j) = x_i \cdot x_j \quad (2.18)$$

Eşitlik (2.19)’da gösterilen polinomial çekirdek fonksiyonu doğrusal çekirdek ile karşılaştırıldığında her bir destek vektörü için uzun hesaplama gerektirir. Doğrusal çekirdek fonksiyonuna göre polinomial çekirdek fonksiyonu daha etkilidir. Polinomial çekirdek fonksiyonunun derecesi genellikle 1’den büyük seçilir. 1 değeri seçilirse polinomial çekirdek fonksiyonu doğrusal çekirdek fonksiyonu olur [39].

$$k(x_i, y_j) = (x_i \cdot x_j + 1)^d \quad (2.19)$$

Karesel çekirdek fonksiyonu kullanılmakta olan çekirdek fonksiyonları arasında en popüler olanıdır. Karesel çekirdek fonksiyonu Eşitlik (2.20)’deki gibi gösterilir [39].

$$k(x_i, y_j) = x_i \cdot x_j (x_i \cdot x_j + 1) \quad (2.20)$$

RBF(Radial basis function) çekirdek fonksiyonu, iyi genelleme performansı sunduğu için geniş kullanım alanı vardır. RBF düzlem ve merkez düzlem destek vektörlerin sayısı ve değerleriyle hesaplanır. Hiper düzlemler arasındaki mesafe minimize edilerek hata payı düşürülebilir. RBF çekirdek fonksiyonu Eşitlik (2.21)’de verilmiştir [40].

$$k(x_i, y_j) = \exp\left(-\gamma |x_i - x_j|^2\right) \quad (2.21)$$

$$\gamma = \left(\frac{1}{2\sigma^2}\right) \quad \gamma > 0 \quad (2.22)$$

γ en yakın veri ile en uygun ayırma düzlemi arasındaki mesafedir. Bu değer σ genişliği değiştirilerek düzenlenebilir.

Pearson VII (PUK) evrensel çekirdek fonksiyonunun kullanılmasının ana nedeni ise w ve σ değerlerinin değiştirilerek esnekliği çok daha ayarlanabilir hale gelmesidir.

Hiper düzlemi oluşturmak için kullanılan The Pearson VII fonksiyonun genel formu [40,41],

$$f(x) = \frac{H}{\left[1 + \left(\frac{2(x-x_0)\sqrt{2\left(\frac{1}{w}\right) - 1}}{\sigma} \right)^2 \right]^w} \quad (2.23)$$

X bağımsız değişken, H x tepe değerleri arasındaki maksimum tepe seviyesi, w ve σ Pearson genişliğini belirtmektedir.

Bu çalışmada, Pearson VII çekirdek fonksiyonu SVM alternatif çekirdek fonksiyonu olarak kullanılmıştır.

Verilerin doğası genellikle bilinmediğinden en uygun kernel fonksiyonunu ve parametrelerini seçmek için deneme yanılma yöntemini kullanarak en yüksek genelleme sonucunu veren performansı seçilir. Yanlış tip seçilen çekirdek fonksiyonu diğer fonksiyonlara göre düşük genelleme sonucu verecektir.

SVM çekirdek fonksiyonlarıyla dönüştürülmüş olan nitelik uzayında, iyi bir genelleme teorisi ve optimizasyon teorisinin yardımıyla doğrusal olarak eğitilen bir öğrenme makinesi olarak adlandırılır.

SVM, öznelik uzayında eğitim verilerine göre en uygun ayırma düzlemini öğrenmeye çalışır [39].

2.3.3.3. LR (Lojistik regresyon) modeli

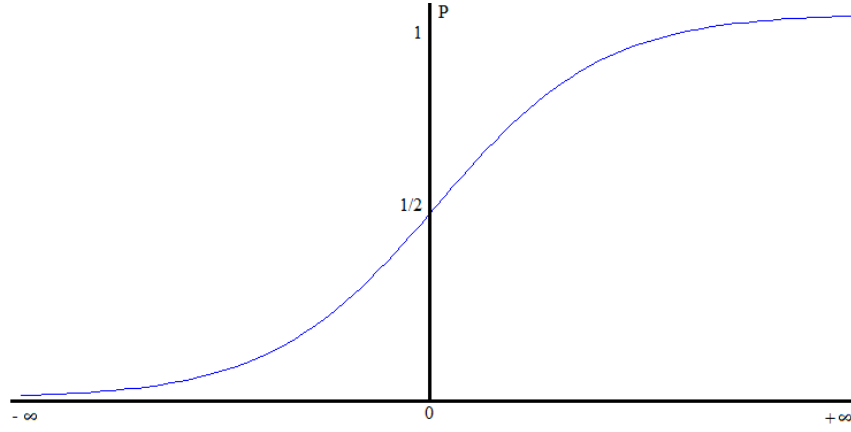
Matematiksel terim olarak regresyon analizi, iki ya da daha çok değişken arasındaki ilişkiyi ölçmek için kullanılan analiz metodudur. Ayrıca regresyon, mevcut eğitim verilerine göre giriş ve çıkış uzayları arasındaki haritalama fonksiyonunu öğrenmeye çalışır [39]. Lojistik regresyon, bağımsız değişkenlerin bağımlı değişkenler üzerinde gösterdiği değişimi olasılık olarak hesaplar, sınıflama yapar.

Eğer tek bir değişken kullanılarak analiz yapılıyorsa buna tek değişkenli regresyon, birden çok değişken kullanılıyorsa çok değişkenli regresyon analizi olarak

isimlendirilir [42]. Lojistik modeli temel alan lojistik fonksiyonun matematiksel formu,

$$F(z) = \frac{1}{1+e^{-z}} \quad (2.24)$$

Lojistik eğrisinin ise, Şekil 2.19'da görüldüğü gibi artan S şeklindedir ve üst sınırı vardır.



Şekil 2.19. Lojistik Regresyon

LR fonksiyonundan yola çıkarak oluşturacağımız lojistik model aşağıda yer almaktadır:

$$Z_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ji} \quad (2.25)$$

X'ler bağımsız değişkenleri göstermektedirler. $B \in \mathbb{R}^d$ maksimum benzerliği hesaplamak için kullanılan regresyon katsayısıdır.

Lojistik regresyon modelinin parametreleri, analitik olarak elde edilemediğinden iteratif bir yöntem olan Maksimum Likelihood (ML) tekniğiyle tahmin edilmektedir. Olabilirlik fonksiyonundaki değişim ihmal edilebilir seviyenin altına indiğinde tekrar etme (iterasyon) işlemine son verilir. LR modeli ortalama π ve varyansı $\pi(1-\pi)$ olan Bernoulli dağılımına uyar. Basit LR modelinde en çok olabilirlik tahmini (Maksimum Likelihood) fonksiyonu Eşitlik (2.26) ve (2.27)'de verilmiştir [43,44]:

$$L = \prod_{i=1}^n \Pr((Y_i=y_i)) \quad (2.26)$$

$$L = \prod_{i=1}^n \frac{[e^{(\beta_0 + \beta_1 x_i)}]^{y_i}}{1 + e^{(\beta_0 + \beta_1 x_i)}} \quad (2.27)$$

$$y_i = 0, 1;$$

2.3.4. Değerlendirme ölçütleri

Sınıflandırıcı algoritmalar kullanılırken verilerin sınıflandırma başarısını değerlendirmek amacıyla farklı teknikler kullanılmaktadır. Bunlar;

- Sınıflama doğruluğu
- Karmaşıklık matrisi (confusion matrix)'dir.

2.3.4.1. Sınıflandırma doğruluğu

Verilerin doğru veya yanlış sınıflandırılma durumları kontrol edilir. Bu duruma göre veri olması gerekenden farklı sınıflandırılmışsa, yanlış olarak değerlendirilir. Olması gereken sınıfta sınıflandırılmışsa, doğru olarak değerlendirilir. Sınıflandırma doğruluğunun başarısı doğru etiketlenilmiş verilerin sayısının, doğru ve yanlış etiketlenilmiş tüm verilerin toplamına oranıdır [28].

2.3.4.2. Karmaşıklık matrisi (Confusion matrix)

Sınıflandırma doğruluğuna benzer şekilde olan bu ölçüt, sadece doğru ve yanlış olarak sınıflandırılma durumlarına bakmak yerine, verilerin gerçek değeri ile sınıflandırma değerlerini de kontrol etmektedir. Doğru sınıflandırılan verilerin yanı sıra yanlış sınıflandırılan verilerin de kontrol edilmesi esasına dayanır.

Tablo 2.2. Karmaşıklık matrisi

		TAHMİNİ SINIFLANDIRMA					TOPLAM EPOK
		NREM1	NREM2	NREM3	REM	WAKE	
GERÇEK SINIFLANDIRMA	NREM1	10	3	2	8	2	25
	NREM2	2	11	1	6	4	24
	NREM3	0	3	18	1	1	23
	REM	8	0	0	14	0	22
	WAKE	2	0	0	0	19	21
TOPLAM EPOK		22	17	21	29	26	

Tablo 2.2’de toplam 125 epok bulunmaktadır. Bunlardan 25 tane NREM1, 24 tane NREM2, 23 tane NREM3, 22 tane REM ve 21 tane WAKE evreleri bulunmaktadır. Sınıflandırıcı 25 tane NREM 1 evresinden 10 tanesini NREM1, 3 tanesini NREM2, 2 tanesini NREM3, 8 tanesini REM, 2 tanesini de WAKE olarak tahmin etmiştir. Doğru tahminlerin hepsi tablonun sol köşeden sağ köşeye doğru giden köşegen hizasında sıralanmıştır [28].

Duyarlılık (Sensitivity) ve özgüllük (Specificity) ikili sınıflandırmada istatistiksel olarak performans ölçme göstergeleridir.

Duyarlılık: Doğru pozitif (True Positive) olarak sınıflandırılmış verileri tespit eder. Yanlış negatif (False Negatif), doğru pozitifin tamamlayıcısıdır. Testin doğru sınıflandırma başarısı ile ilgili orandır.

$$\text{Duyarlılık} = \frac{\text{Doğru Pozitif}}{\text{Doğru pozitif} + \text{Yanlış Negatif}} \quad (2.28)$$

Özgüllük: Doğru negatif olarak da kullanılabilir. Kendi sınıfından olmayan verilerin, kendi sınıfı dışındaki sınıflarda sınıflandırılmasıyla, doğru olarak sınıflandırılmasıdır.

Diğer bir deyişle, doğru olarak sınıflandırılmış negatiflerin oranıdır. Yanlış pozitif oranının tamamlayıcısıdır.

$$\text{Özgüllük} = \frac{\text{Doğru Negatif}}{\text{Doğru Negatif} + \text{Yanlış Pozitif}} \quad (2.29)$$

Mükemmel bir sınıflandırıcı %100 duyarlılık ve özgüllük oranına sahiptir. Genel doğruluk, hasta ve sağlam kişilerin yüzde kaçının önerilen yöntemle tanınabildiğini gösterir. Doğru sınıflandırılmış örnek sayısının, tüm örnek sayısına bölünmesidir. Hata ise yanlış sınıflandırılmış örnek sayısının, tüm örnek sayısına bölünmesidir.

$$\text{Doğruluk} = \frac{\text{DP} + \text{DN}}{\text{DP} + \text{DN} + \text{YN} + \text{YP}} \quad (2.30)$$

NREM1 evresi için doğru pozitif, yanlış pozitif, yanlış negatif, doğru negatif kavramlarını örnekle açıklanırsa,

Doğru pozitif, doğru sonuçlandırılan tahmini veri sayısıdır. Tablo 2.3’de 25 adet NREM1 evresine ait epoktan 10 tanesinin doğru tahmin edildiği görülmektedir.

Tablo 2.3. Doğru Pozitif

		TAHMİNİ SINIFLANDIRMA				
		NREM1	NREM2	NREM3	REM	WAKE
GERÇEK SINIFLANDIRMA	NREM 1	10	3	2	8	2
	NREM 2	2	11	1	6	4
	NREM 3	0	3	18	1	1
	REM	8	0	0	14	0
	WAKE	2	0	0	0	19

Tablo 2.4’de NREM1 olmayan 12 tane epoğun NREM1 olarak tahmin edilmiş olduğu görülmektedir.

Tablo 2.4. Yanlış pozitif

		TAHMİNİ SINIFLANDIRMA				
		NREM1	NREM2	NREM3	REM	WAKE
GERÇEK SINIFLANDIRMA	NREM 1	10	3	2	8	2
	NREM 2	2	11	1	6	4
	NREM 3	0	3	18	1	1
	REM	8	0	0	14	0
	WAKE	2	0	0	0	19

Tablo 2.5’de NREM1 olan 15 adet epok NREM1 sınıflandırılmayarak yanlış skorlanmıştır.

Tablo 2.5. Yanlış negatif

		TAHMİNİ SINIFLANDIRMA				
		NREM1	NREM2	NREM3	REM	WAKE
GERÇEK SINIFLANDIRMA	NREM 1	10	3	2	8	2
	NREM 2	2	11	1	6	4
	NREM 3	0	3	18	1	1
	REM	8	0	0	14	0
	WAKE	2	0	0	0	19

Tablo 2.6’da, NREM1 olmayan 78 adet epok NREM1 sınıflandırılmayarak doğru sınıflandırılmış olduğu görülmektedir.

Tablo 2.6. Doğru negatif

		TAHMİNİ SINIFLANDIRMA				
		NREM1	NREM2	NREM3	REM	WAKE
GERÇEK SINIFLANDIRMA	NREM 1	10	3	2	8	2
	NREM 2	2	11	1	6	4
	NREM 3	0	3	18	1	1
	REM	8	0	0	14	0
	WAKE	2	0	0	0	19

Tablo 2.7. Karmaşıklık Matrisi

Doğru Pozitif 10	Yanlış Pozitif 14
Yanlış Negatif 15	Doğru Negatif 78

Tablo 2.7’deki verileri Eşitlik (2.30)’daki eşitlikte yerine koyulursa 0,78 değeri elde edilir. NREM1 evresi için başarı oranının %78 olduğu söylenebilir.

2.4. Uygulamada Kullanılan Veri Madenciliği Yazılımı

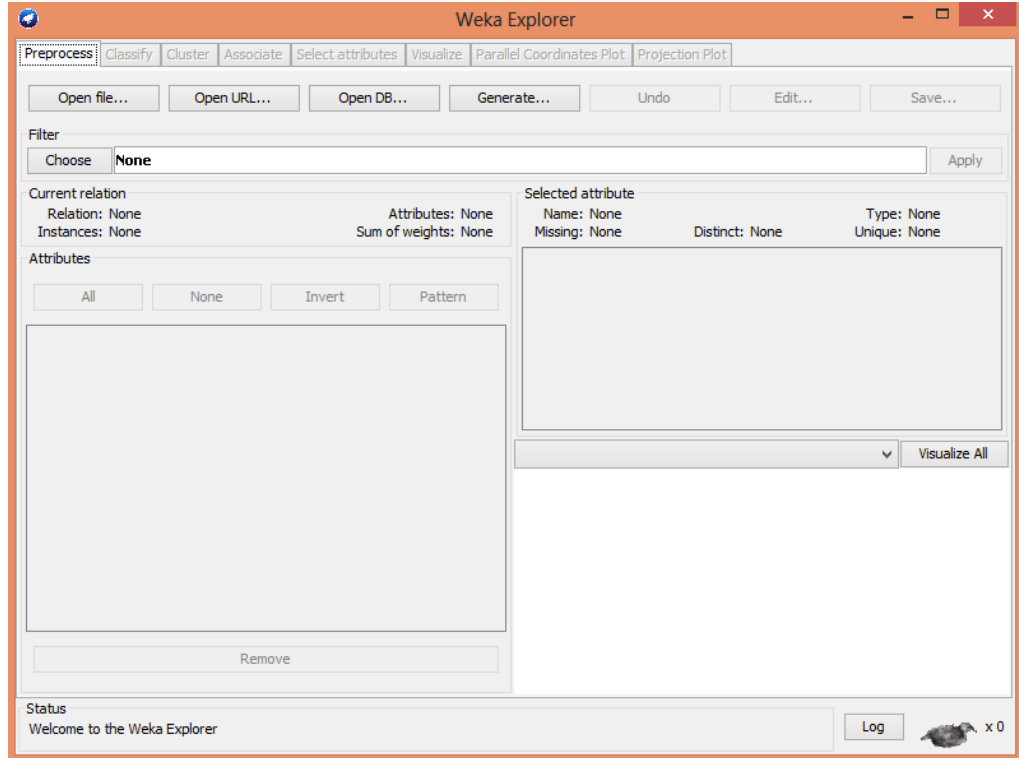
2.4.1. Weka

Waikato Environment for Knowledge Analysis Waikato Üniversitesinin internet sitesinden yüklenebilen java destekli bir veri madenciliği programıdır [28]. Şekil 2.2’de ara yüzü görünen Weka’nın, veri madenciliği problemlerini çözmek için geliştirilen tüm algoritmaların bir arada sunması ve kolay ulaşılabilirliği en önemli avantajlarıdır [32].

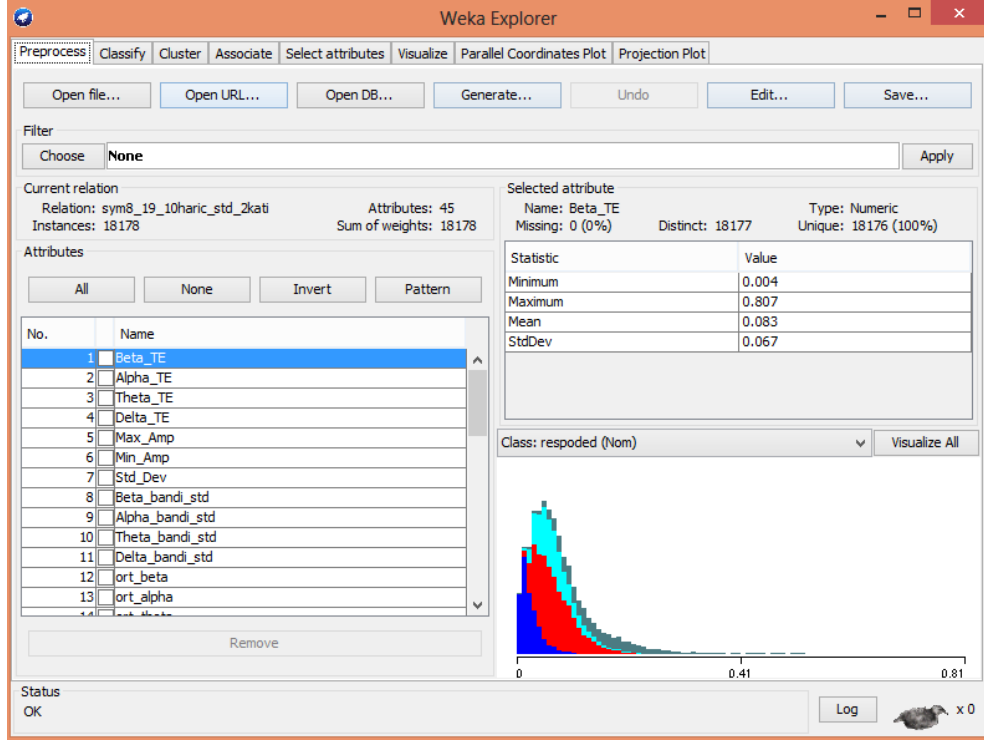


Şekil 2.20. Weka Genel Ara yüz görünümü

Matlab ortamından elde edilen veriler Weka programına aktarılmıştır. Şekil 2.20’de Weka programı ara yüz görüntüsü gösterilmiştir. Explorer sekmesi tıklanıldığında Şekil 2.21’deki görüntü elde edilir. Veriler Weka programına yüklendikten sonra elde edilen ara yüz görüntüsü Şekil 2.22’de gösterilmektedir.

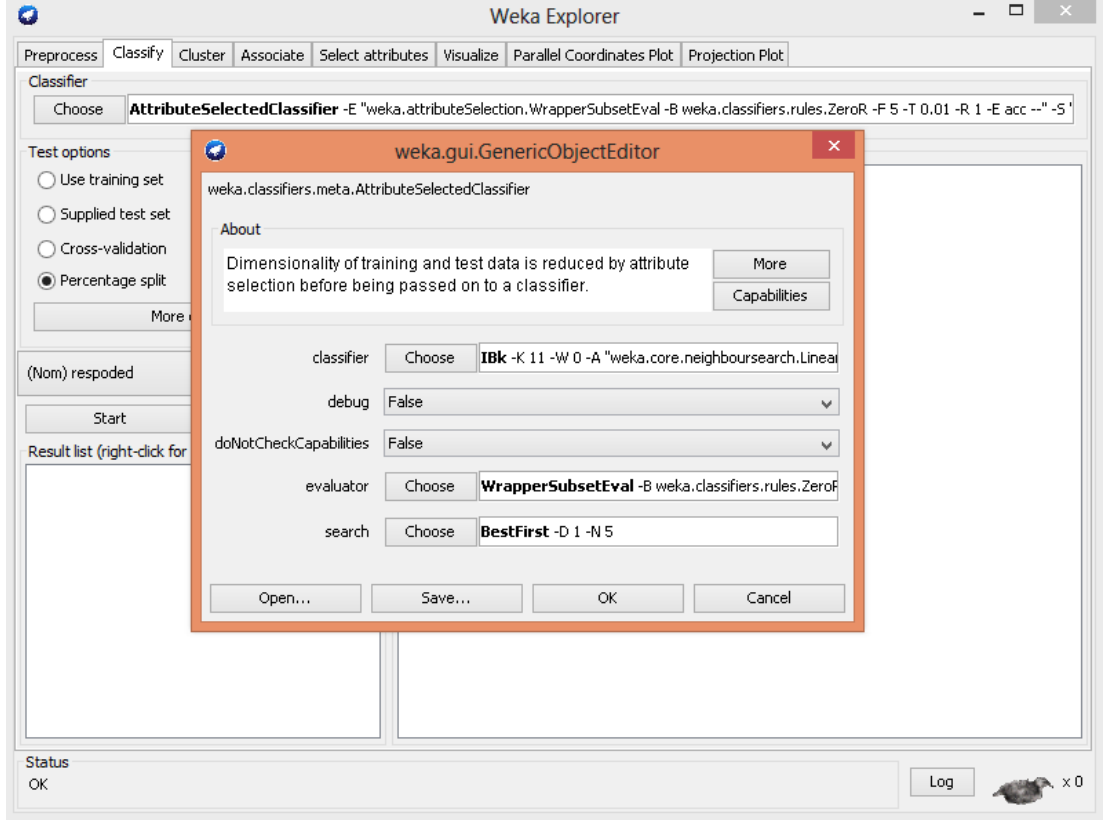


Şekil 2.21. Weka veriler yüklü değil iken ara yüz görünümü



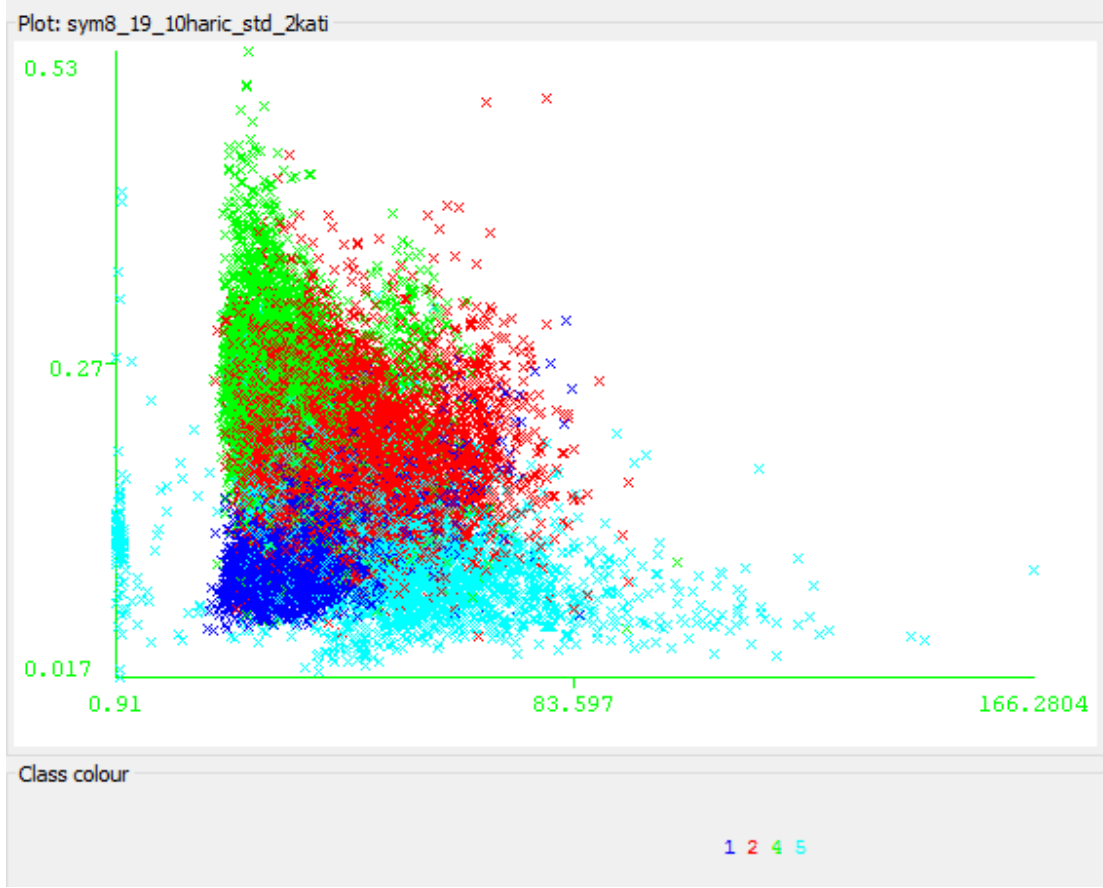
Şekil 2.22. Weka veriler yüklü iken elde edilen ara yüz görünümü

Bu çalışmada öznitelikler sınıflandırıcıya uygulanmadan önce öznitelik azaltımı gerçekleştirilmiştir. Öznitelik azaltımının otomatik olarak sınıflandırıcıya uygulanması amacıyla, meta temelli Attirube Selected Classifier seçilerek uygun sınıflandırıcı ve özellik azaltımı için gerekli olan bilgiler Weka programına girilmiştir. Attirube Selected Classifier, eğitim ve test verilerini bir sınıflandırıcıya uygulanmadan önce öznitelik azaltımını gerçekleştirir. Sonrasında kullanıcıya sorulmadan önceden bilgileri girilmiş olan sınıflandırıcıya uygular. Şekil 2.23'de özellik azaltımı ve sınıflandırıcı ara yüz ekran görüntüsü gösterilmektedir.



Şekil 2.23. Weka özellik azaltımı

Şekil 2.24’de delta enerji değerinin teta enerjisi/toplam enerji değerine oranlanması ile 14542 adet epöğün uyku evresi sınıflarına dağılımının grafiği gösterilmektedir. Grafikten bu oranlama sonucunda uyanık evresini diğer evrelere göre daha iyi ayırt edilebildiği, NREM3 evresi ve NREM2 evresi için bu oranın başarılı bir ayırım yapamadığı söylenebilir [45].



Şekil 2.24. Epokların uyku evresi sınıflarına dağılımı

Şekil 2.25. Eğitim ve test verisi oranı ekran görünümü

Şekil 2.25 verilerin ne kadarının eğitim ne kadarının test verisi olarak kullanılacağını belirleyen ekran görüntüsüdür. Buna göre yüklenen verilerin %20'si eğitim aşaması için %80'i ise test aşaması için kullanılır.

2.4.1.1. Weka da kullanılan özellikler

Veri ön işleme: Weka CSV, JSON, Matlab ASCII ve kendi uzantısı olan Arff veri formatlarını desteklemektedir.

Veri Sınıflandırma: Büyük veriler üzerinde bayes algoritmaları, destek vektör makinaları gibi matematiksel fonksiyonlar, meta temelli algoritmalar kullanılarak sınıflandırma yapılabilmektedir.

Veri kümeleme: Kümeleme için k-ortalama yöntemi algoritmaları yanı sıra yoğunluk ve hiyerarşik kümeleme algoritmalarını da kullanmak mümkündür.

Veri Görselleştirme: Veride kullanılan özniteliklerin birbirlerine göre oranları gibi bir çok grafiği görsel olarak inceleme imkânı mevcuttur [46].

2.4.1.2. Weka dosya formatı

Wekanın kullandığı dosya formatı Arff (Attribute Relationship File Format) dir. Metin dosyası düzenleyici ile düzenlenebilir. Yapısı şekil 2.26'daki gibidir.

```
% Yorum satırı
@RELATION ***** % Veri seti ismi
@ATTRIBUTE ***** numeric % Sayısal Değer
@ATTRIBUTE ***** nominal % karakter
@ATTRIBUTE ***** {1,2,4,5 } % Sonuçlar
@DATA % Eldeki öznitelikler
0.22278713,0.303764887,1
0.265570701,0.396487506,1
```

Şekil 2.26. Weka genel dosya formatı

Data bölümü altında yer alan her bir satır bir epoğa aittir. Bu örnekte iki öznitelik ve sınıflandırma sonuçlarına yer verilmiştir. Attiribute bölümünde kullanılacak olan özniteliklerin her bir satıra ayrı ayrı eklenmesi gerekmektedir. Relation kısmında ise veri setine genel bir isim verilmelidir. Bu çalışmada kullanılan dosya formatı yapısı şekil 2.27'deki gibidir.

```

@RELATION db44
@ATTRIBUTE Beta_TE NUMERIC
@ATTRIBUTE Alpha_TE NUMERIC
@ATTRIBUTE Theta_TE NUMERIC
@ATTRIBUTE Delta_TE NUMERIC
@ATTRIBUTE Max_Amp NUMERIC
@ATTRIBUTE Min_Amp NUMERIC
@ATTRIBUTE Std_Dev NUMERIC
@ATTRIBUTE Beta_bandi_std NUMERIC
@ATTRIBUTE Alpha_bandi_std NUMERIC
@ATTRIBUTE Theta_bandi_std NUMERIC
@ATTRIBUTE Delta_bandi_std NUMERIC
@ATTRIBUTE ort_beta NUMERIC
@ATTRIBUTE ort_alpha NUMERIC
@ATTRIBUTE ort_theta NUMERIC
@ATTRIBUTE ort_delta NUMERIC
@ATTRIBUTE max_beta NUMERIC
@ATTRIBUTE min_beta NUMERIC
.
.
.
@ATTRIBUTE respoded {1,2,4,5}
@DATA
0.180330285,0.348111128,0.1039028,0.366421989,0.095228219,-
0.047614109,0.029335293,0.01066795,0.017310507,0.009457237,0.017753429,
6.5304E-07,-5.27252E-07,8.54216E-06,0.000480961,0.046442886,-
0.047776111,0.069001203,-0.067313,0.033638085,-0.040141656,0.084178279,-
0.048428934,50.59077366,79.90735947,43.77024058,82.58674399,-
5.677311613,-13.06415,-4.54635437,-
13.57239182,6.164249499,1.297131481,3.699333325,1.657079347,0.299366688,
0.036421679,0.000128489,0.295565643,0.03005283,2.90732E-
06,0.1153983,0.580994581,0.339703253,5

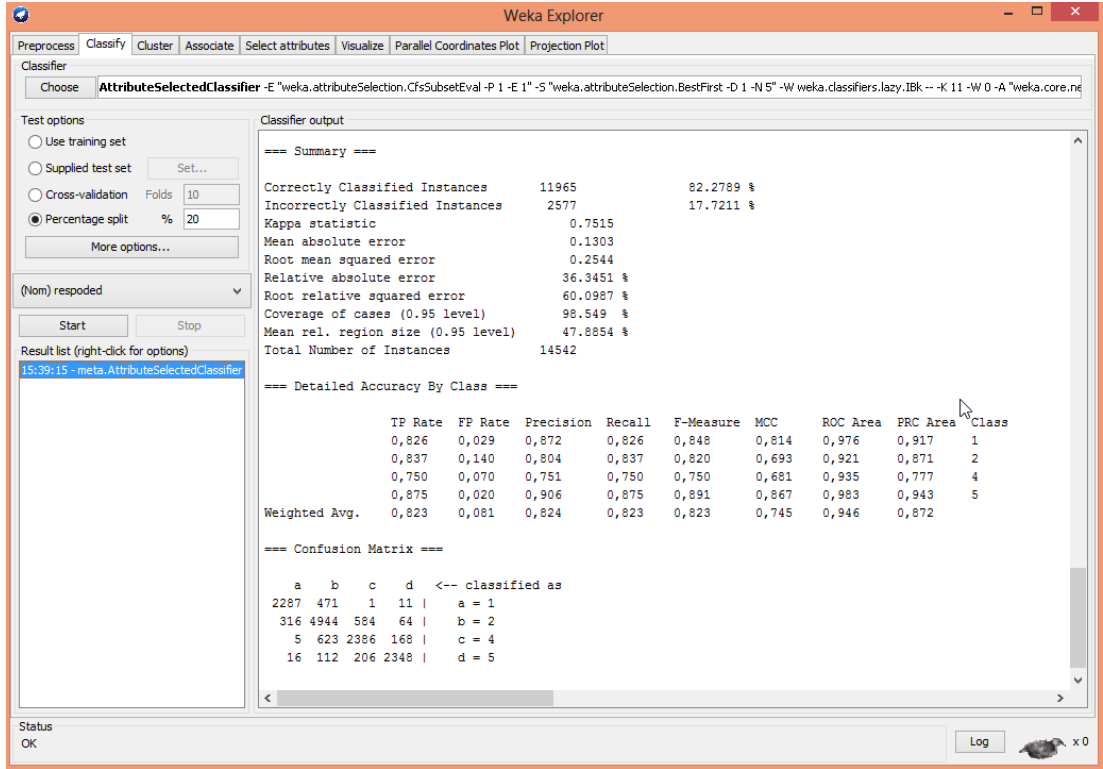
```

Şekil 2.27. Çalışmada kullanılan weka dosya formatı

2.4.1.3. Weka sonuç ekranı

Şekil 2.28’de test verisi olarak kullanılan %80 oranındaki veri setinin sınıflandırma sonucu gösterilmiştir. Bu şekildeki,

Correctly classified instances sınıfı doğru tespit edilen verinin sayısını, Incorrectly Classified Instances sınıfı yanlış belirlenen verinin sayısını, Mean absolute error ise tam hata değeri göstermektedir. Total Number of Instances toplam örnek sayısı değerini vermektedir.



Şekil 2.28. Weka sonuç ekranı

3. BULGULAR

Bu çalışmada, dreamspjject veritabanında yer alan veriler kullanılmıştır. Bu veriler 18 hastadan Belçika hastanesi uyku laboratuvarından 32 kanallı PSG cihazı ile en az iki EOG kanalı, 3 EEG kanalı ve bir tane EMG kanalından kayıt alınmıştır. Veriler edf formatında kaydedilmiştir. NREM1 ve REM evrelerinin ayırt edici özellikleri bir birlerine çok benzediğinden dolayı bu iki evre tek evre olarak göz önünde bulundurulmuştur. Verinin örnekleme frekansı 200 Hz'dir.

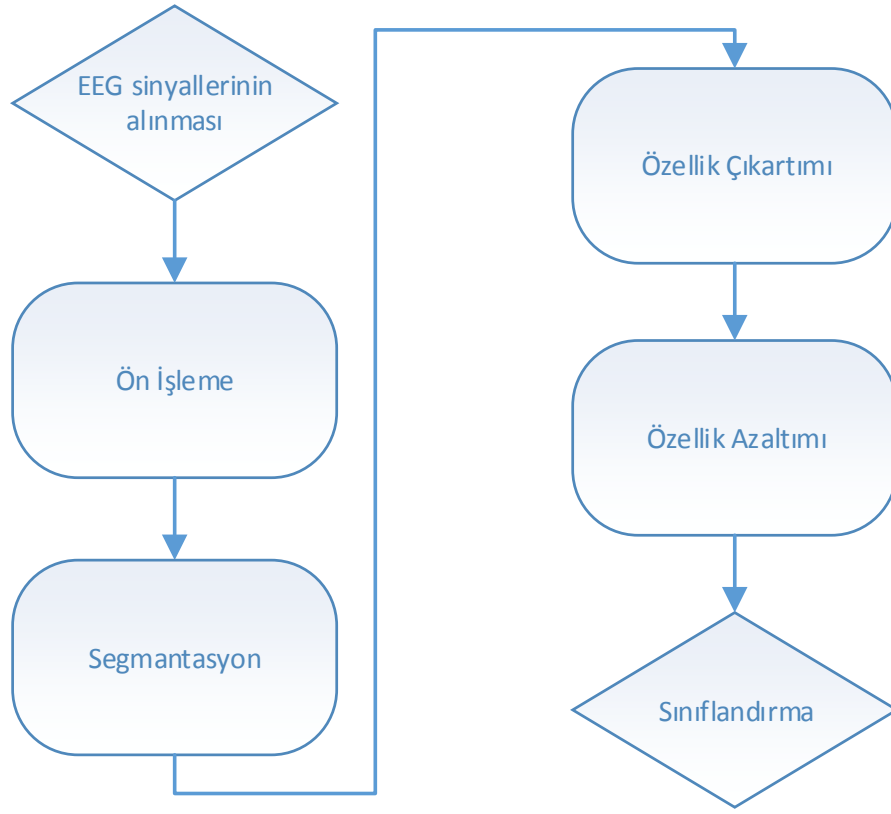
Uyku evreleri Rechtschaffen and Kales kriterlerine göre 20 sn'lik epoklar halinde, AASM kriterlerine göre ise 30 sn'lik epoklar halinde skorlanmıştır. Her iki skorlama da uyku laboratuvarında uzman tarafından gerçekleştirilmiştir. Bu veriler dreamspjject veritabanında metin dosyası olarak yer almaktadır. Elde edilen bu veriler uyku evreleri skorlamasında referans alınmıştır ve doğruluk oranları bu verilere dayanarak tespit edilmiştir [47].

Tablo 3.1'de uyku evresi skorlaması yapılan hastalara ait istatistiksel veriler bulunmaktadır.

Tablo 3.1. Uyku evresi skorlaması yapılan hastalara ait istatistiksel veriler

AD	ÖRNEKLEME FREKANSI	YAŞ	CİNSİYET	KAYIT SÜRESİ
Subject1	200 Hz	23	Bayan	08:00:40
Subject2	200 Hz	47	Bayan	08:12:30
Subject3	200 Hz	24	Bayan	08:24:20
Subject4	200 Hz	48	Bayan	08:46:30
Subject5	200 Hz	46	Bayan	08:51:30
Subject6	200 Hz	65	Bayan	08:18:40
Subject7	200 Hz	45	Bayan	08:26:00
Subject8	200 Hz	22	Bayan	08:05:00
Subject9	200 Hz	21	Bayan	09:18:40
Subject11	200 Hz	30	Bayan	08:24:10
Subject12	200 Hz	54	Bayan	08:00:40
Subject13	200 Hz	23	Bayan	09:15:40
Subject14	200 Hz	57	Bayan	08:22:10
Subject15	200 Hz	20	Bayan	07:00:00
Subject16	200 Hz	27	Bayan	08:03:50
Subject17	200 Hz	23	Bay	08:18:30
Subject18	200 Hz	27	Bay	08:30:40
Subject20	200 Hz	20	Bay	09:32:30

Tablo 3.1’den görüldüğü üzere bu çalışmada 18 hastaya ait ortalama 8 saat süre ile çekilmiş olan veriler kullanılmıştır. Şekil 3.1’de bu çalışmada gerçekleştirilen sistemin akış diyagramı verilmiştir.



Şekil 3.1. Önerilen sistemin akış diyagramı

3.1. EEG - EOG Sinyallerinin Alınması

Matlab’da yazılmış olan edfread adlı fonksiyon ile edf dosyasınının 23 farklı kanaldan alınan header ve record kayıtları elde edilmiştir. Bu çalışmada CZA1, EOG1 ve EOG2 kanallarından alınan veriler kullanılmıştır. Uzman skorlamaları dreamsprojet veritabanında *.txt dosyası olarak verilmiş bu verilerde Matlab ortamına alınmıştır. 18 hastadan ayrı ayrı alınan PSG kayıtlarınının toplam süresi 151 saat 52 dakikadır.

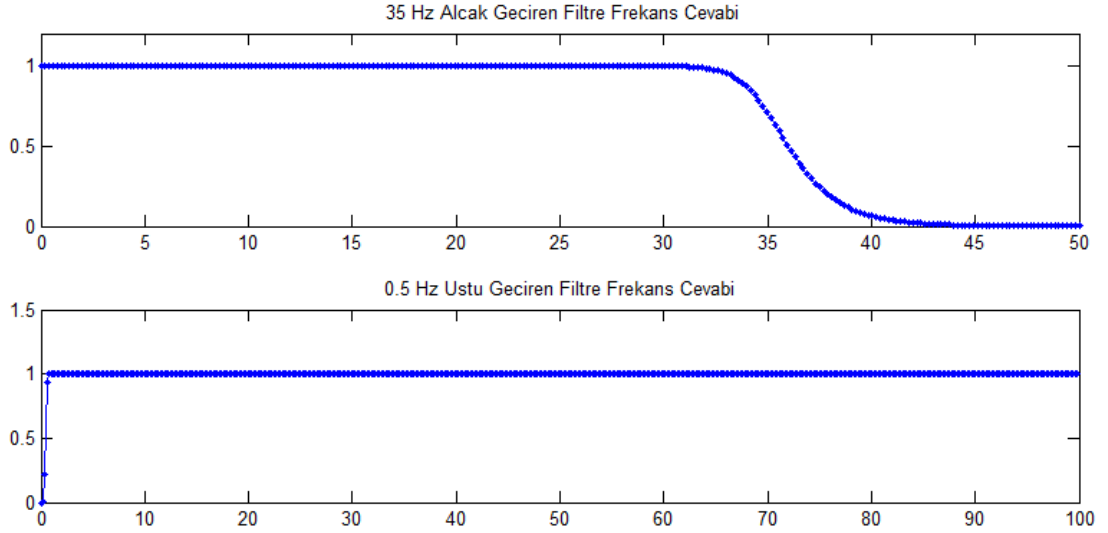
Edf uzantılı veri dosyasınının Matlab ortamında okunmasıyla record ve header adlı iki başlık elde edilmiştir. Record başlığı altında hastalardan alınmış sinyaller, header başlığı altında ise bu kayıtlara ait çeşitli bilgiler kullanıma hazır hale gelmiştir. Şekil 3.2’de edf uzantılı dosyanın Matlab ortamına alınmasıyla elde edilen header dosyası içerisindeki başlıklar görülmektedir. Burada PSG cihazı aracılığıyla alınan kayıtlarla ilgili gereken bilgileri incelemek mümkün olmaktadır.

Field	Value	Min	Max
ver	0	0	0
patientID	'DatabaseSubjects, su...		
recordID	'STARTDATE : 29/05/...		
startdate	'29.05.02'		
starttime	'22.55.00'		
bytes	6144	6144	6144
records	14420	14420	14420
duration	2	2	2
ns	23	23	23
label	<1x23 cell>		
transducer	<1x23 cell>		
units	<1x23 cell>		
physicalMin	<1x23 double>	-8000	40
physicalMax	<1x23 double>	3	8000
digitalMin	<1x23 double>	-5914	299
digitalMax	<1x23 double>	473	5915
prefilter	<1x23 cell>		
samples	<1x23 double>	400	400

Şekil 3.2. Edf uzantılı verinin Matlab ortamına aktarılması

3.2. Ön İşleme

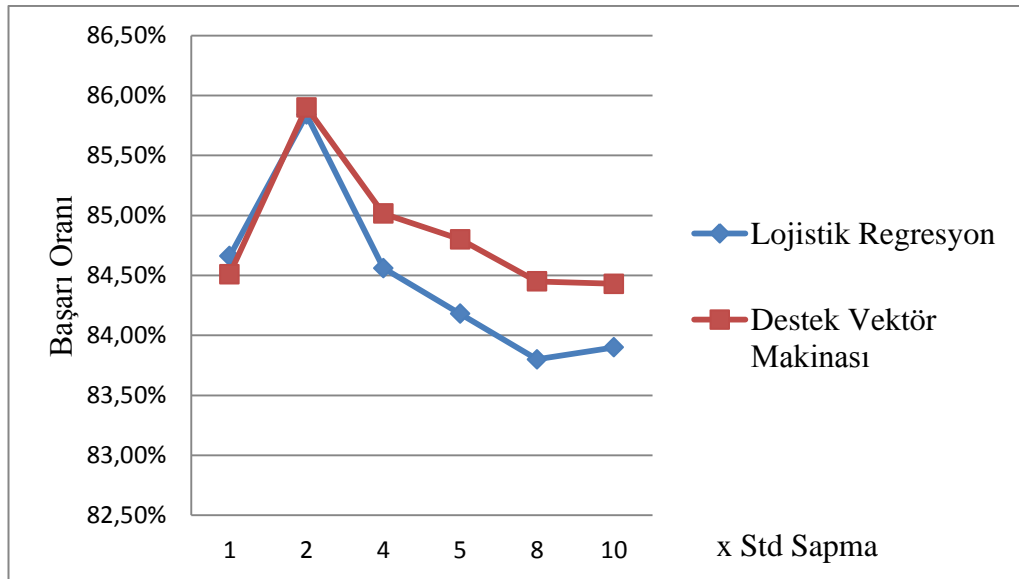
CzA1 kanalından alınan sinyaller ilk olarak bir alçak geçiren filtreden geçirilir. Sinyalin filtrelenmesinde 16. dereceden 35 Hz ve alt frekansları geçiren Alçak geçiren Butterworth IIR(Infinite Impulse Response) dijital filtre kullanılmıştır. Elde edilen filtrelenmiş sinyal 6.dereceden 0,5 Hz ve üstünü geçiren yüksek geçiren Butterworth IIR dijital filtreden geçirilmiştir. Bant geçiren filtre yerine ayrı ayrı filtre kullanılmasının nedeni, alçak geçiren filtrenin frekans cevabında görüldüğü gibi daha yüksek dereceden filtre tasarlanıp kesim frekans eğimini daha dik elde edebilmektir. 0,5 – 35 Hz aralığında elde edilen sinyal uyku EEG sinyallerinin bulunduğu frekans bandıdır [11]. EOG1 ve EOG2 kanallarından alınan veriler ise 8.dereceden 5 Hz alçak geçiren ve 5-35 Hz bant geçiren Butterworth IIR dijital filtreden geçirilerek iki ayrı frekans bandında incelenmiştir.



Şekil 3.3. Filtre Cevabı

Ham PSG verileri filtrelendikten sonra CZA1 kanalından elde edilen veriler, standart sapmanın 2 katından büyük veya küçük olan değerler için standart sapmanın 2 katına sabitlenmiştir.

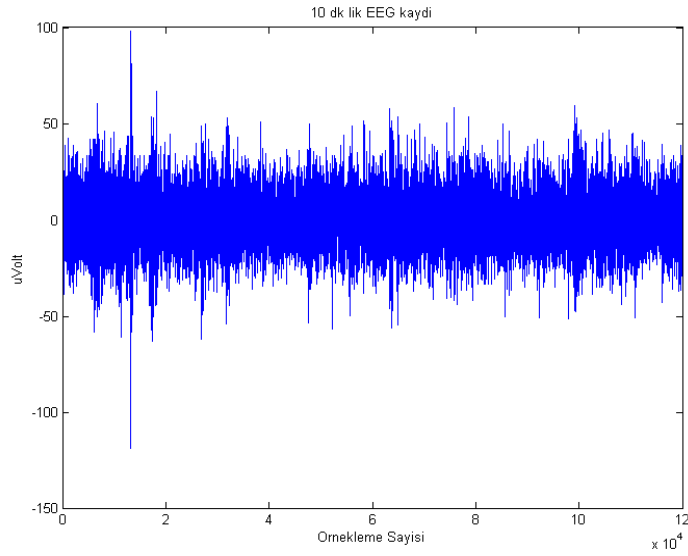
Lojistik regresyon ve SVM kullanılarak yapılan sınıflandırmada standart sapmanın sınıflandırma sonucuna olan etkisi aşağıdaki Şekil 3.4’de gösterilmiştir. Bu şekil sonucunda standart sapmanın iki katı kadar olan verilerin seçilmesi durumunda başarının arttığı gözlemlenmiştir.



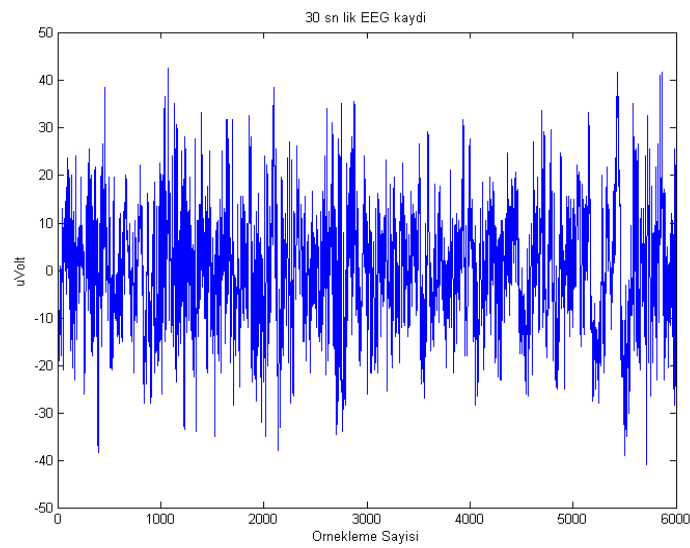
Şekil 3.4. Standart sapmanın sabitlenmesiyle sınıflandırıcı başarısının değişimi

3.3. Segmantasyon

Segmantasyon PSG cihazından elde edilen verilerin (CZA1, EOG1, EOG2 kanalı) uyku evresi skorlaması için gerekli olan 30 sn'lik sürelerle bölünmesidir. AASM standartlarına göre en fazla 30 sn'den oluşan epoklar skorlanarak uyku evresi skorlaması elde edilmektedir [10]. Elde edilen her bir 30 sn'lik bölüme bir epok denir. Bu çalışmada, segmentasyon için Matlab'da fonksiyon yazılmıştır. Bu segmentasyon sonucunda 18 hastadan 14542 adet epok elde edilmiş ve sınıflandırılmıştır. Şekil 3.5 ve 3.6'da örnek EEG kayıtları verilmiştir.



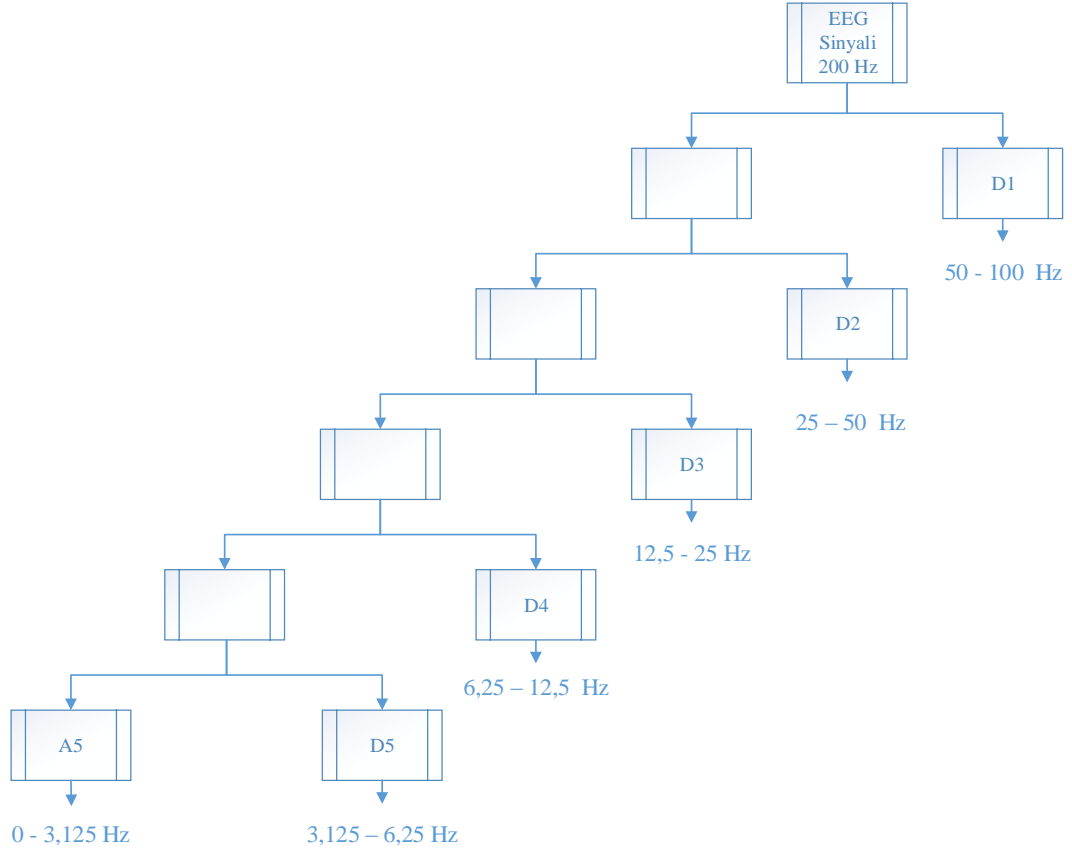
Şekil 3.5. 10 dk'lık EEG kaydı



Şekil 3.6. 30 sn'lik EEG kaydı

3.4. Öznitelik Çıkartımı

Uyku evresi skorlaması için ayrık dalgacık yöntemiyle delta, teta, alfa ve beta bandı frekans aralıklarını elde edebilmek için 4 alt banda ayrıştırılmıştır. Her bir seviye örnekleme frekansının yarısı oranında ayrıştırma yapmaktadır. Şekil 3.7’de bu ayrışım seviyeleri gösterilmiştir.



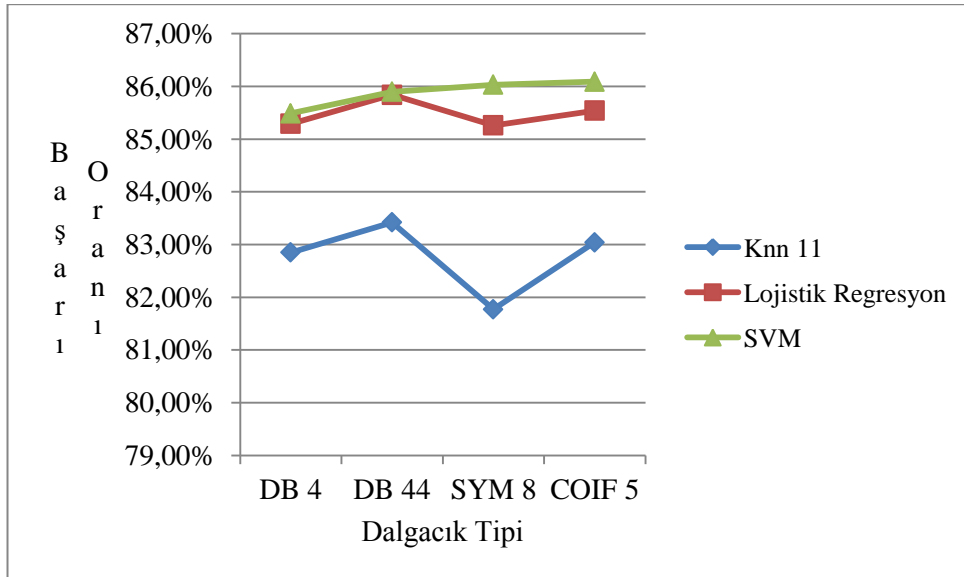
Şekil 3.7. Ayrışım seviyeleri

Bu ayrışım seviyelerinin karşılık geldiği baskın uyku evresi ve frekans aralıkları dalgacık katsayılarıyla beraber Tablo 3.2’de gösterilmiştir.

Tablo 3.2. EEG sinyali alt frekans bantları

Dalga Tipi	Dalgacık katsayıları	Frekans Aralığı (Hz)	Baskın Görüldüğü Evre
Delta	A5	0 – 3,125	NREM3
Teta	D5	3,125 – 6,25	NREM1
Alfa	D4	6,25 – 12,5	Gözler Kapalı(Pasif) Uyanık
Beta	D3+D2	12,5 – 35	Gözler Açık (Aktif)Uyanık

Frekans domeninde yapılan ayrıştırma için 4 farklı dalgacık tipinde simülasyon yapılmıştır. Şekil 3.8’de 4 farklı dalgacık tipinin 3 farklı sınıflandırıcıda öznelik azaltımı yapılmadan göstermiş oldukları başarı oranları gösterilmektedir.



Şekil 3.8. Öznelik azaltımı yapılmadan dalgacık tipine göre başarı oranları

Şekil 3.8’de SVM sınıflandırıcısının Sym 8 ve Coif 5 dalgacığı ile yapılan ayrıştımda daha başarılı olduğu, Db 44 dalgacığının ise KNN 11 ve Lojistik regresyon sınıflandırıcılarında başarılı olduğu, Db 4 dalgacığının diğer dalgacık çeşitlerine göre daha az başarı oranı göstermiş olduğu gözlemlenmiştir.

Elde edilen her bir ayrışım seviyesi kullanılarak her bir epok için öznitelik çıkartımı yapılmış ve Matlab programı aracılığıyla hesaplatılmıştır. Tablo 3.3’de bir epok için elde edilen öznitelikler gösterilmiştir.

Tablo 3.3. Bir epok için elde edilen öznitelikler

Sıra No	Öznitelikler	Sıra No	Öznitelikler
1	Beta Dalgası Enerjisi / Toplam Enerji	2	Alfa Dalgası Enerjisi / Toplam Enerji
3	Teta Dalgası Enerjisi / Toplam Enerji	4	Delta Dalgası Enerjisi / Toplam Enerji
5	Epoğun maksimum genliği	6	Epoğun minimum genliği
7	Epoğun Standart Sapması	8	Beta bandının standart sapması
9	Alfa bandının standart sapması	10	Teta bandının standart sapması
11	Delta bandının standart sapması	12	Beta bandının Ortalaması
13	Alfa bandının Ortalaması	14	Teta bandının Ortalaması
15	Delta bandının Ortalaması	16	Beta bandının maksimum genlik değeri
17	Beta bandının minimum genlik değeri	18	Alfa bandının maksimum genlik değeri
19	Alfa bandının minimum genlik değeri	20	Teta bandının maksimum genlik değeri
21	Teta bandının minimum genlik değeri	22	Delta bandının maksimum genlik değeri
23	Delta bandının Minimum genlik değeri	24	Beta bandı enerjisi
25	Alfa bandı enerjisi	26	Teta bandı enerjisi
27	Delta bandı enerjisi	28	Beta bandı entropi değeri
29	Alfa bandı entropi değeri	30	Teta bandı entropi değeri
31	Delta bandı entropi değeri	32	5 Hz altı EOG1 sinyali Enerjisi
33	5 Hz üstü EOG1 sinyali Enerjisi	34	5 Hz altı EOG2 sinyali Enerjisi

Tablo 3.3.(Devam) Bir epok için elde edilen öznitelikler

35	5 Hz üstü EOG2 sinyali Enerjisi	36	1.kanal EOG'sinin tepeden tepeye genliği
37	1.kanal EOG'sinin standart sapması	38	1.kanal EOG'sinin ortalaması
39	2.kanal EOG'sinin tepeden tepeye genliği	40	2.kanal EOG'sinin standart sapması
41	2.kanal EOG'sinin ortalaması	42	Hjorth parametresi 1(activity)
43	Hjorth parametresi 2(mobility)	44	Hjorth parametresi 3(complexity)

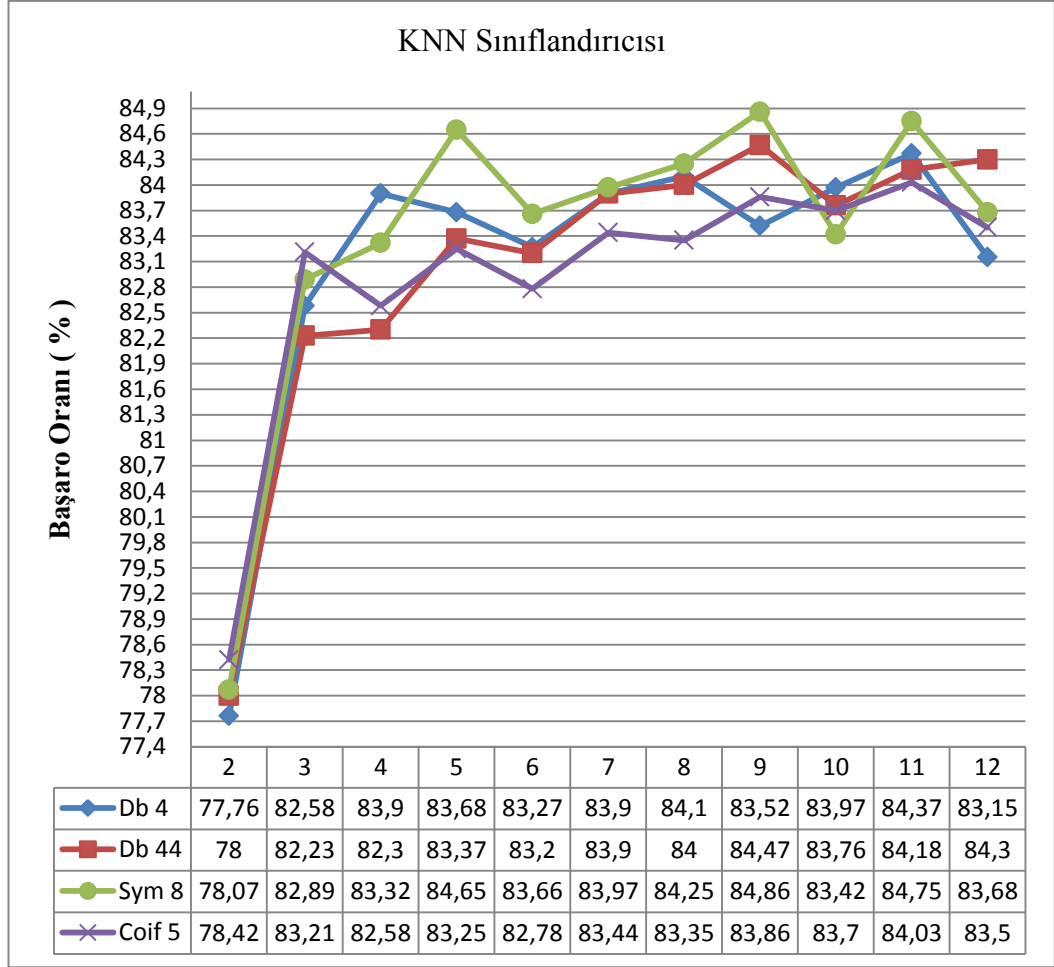
3.5. Öznitelik Azaltımı

Elde edilen öznitelikler sınıflandırmadaki ayrımı iyi yapamadığında, sınıflandırıcılar da istenilen performansı gösteremeyebilir. Bundan dolayı sınıflandırılacak veriyi iyi temsil edebilecek öznitelikler seçilmelidir. Öznitelik seçme ve azaltma yöntemleri; öznitelik havuzunda bulunan ilgisiz öznitelikleri havuzdan çıkartmak ve verileri daha kolay ayrıştırılabilir hale getirmek için kullanılmaktadır. Bu sayede, sınıflandırıcının performansı artmakta aynı zamanda gereksiz işlem yükü de azaltılmış olmaktadır.

Matlab programı aracılığıyla EEG sinyallerinden elde edilen öznitelikler, Weka version 3.7.11 programına aktarılarak öznitelik azaltımı gerçekleştirilmiştir. İki farklı özellik değerlendiricisi aracıyla ayrı ayrı sonuçlar gözlemlenmiştir.

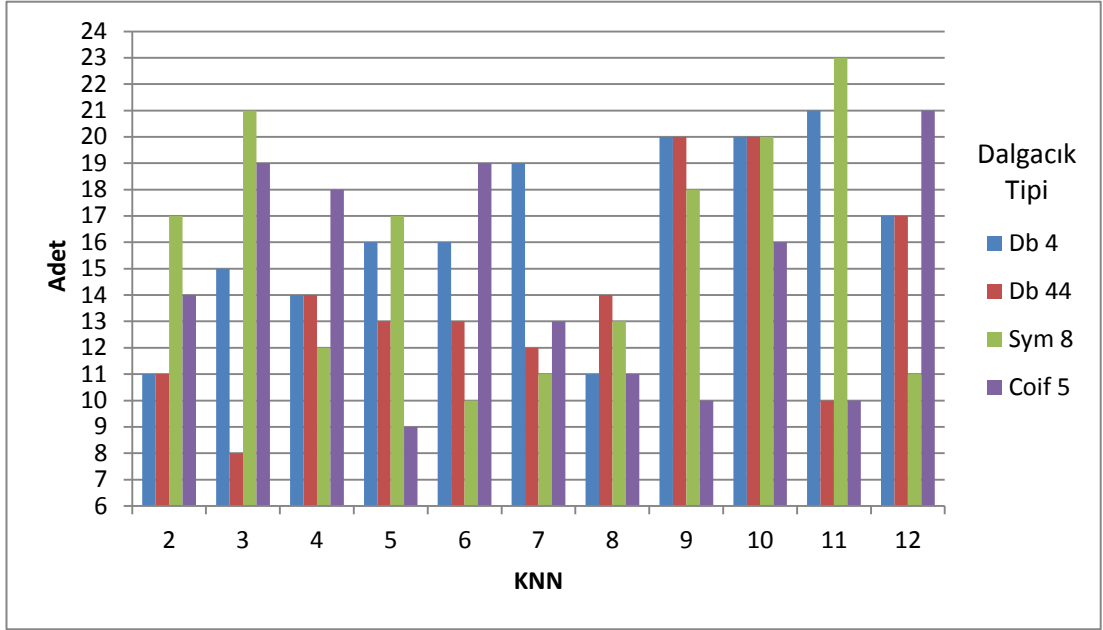
3.5.1. Sarmal metot ile KNN sınıflandırıcısı

Şekil 3.9'da KNN sınıflandırıcısı K değerine göre sarmal metot ile öznitelik azaltımı sonucunda elde edilmiş olan başarı oranları görülmektedir. En yüksek başarı oranının %84,86 ile K=9 değerinde Sym 8 dalgacığından elde edilmiş ve Şekil 3.9'da elde edilen sonuçlar gösterilmiştir.



Şekil 3.9. Dalgacık tipine göre KNN sınıflandırıcısı başarısı

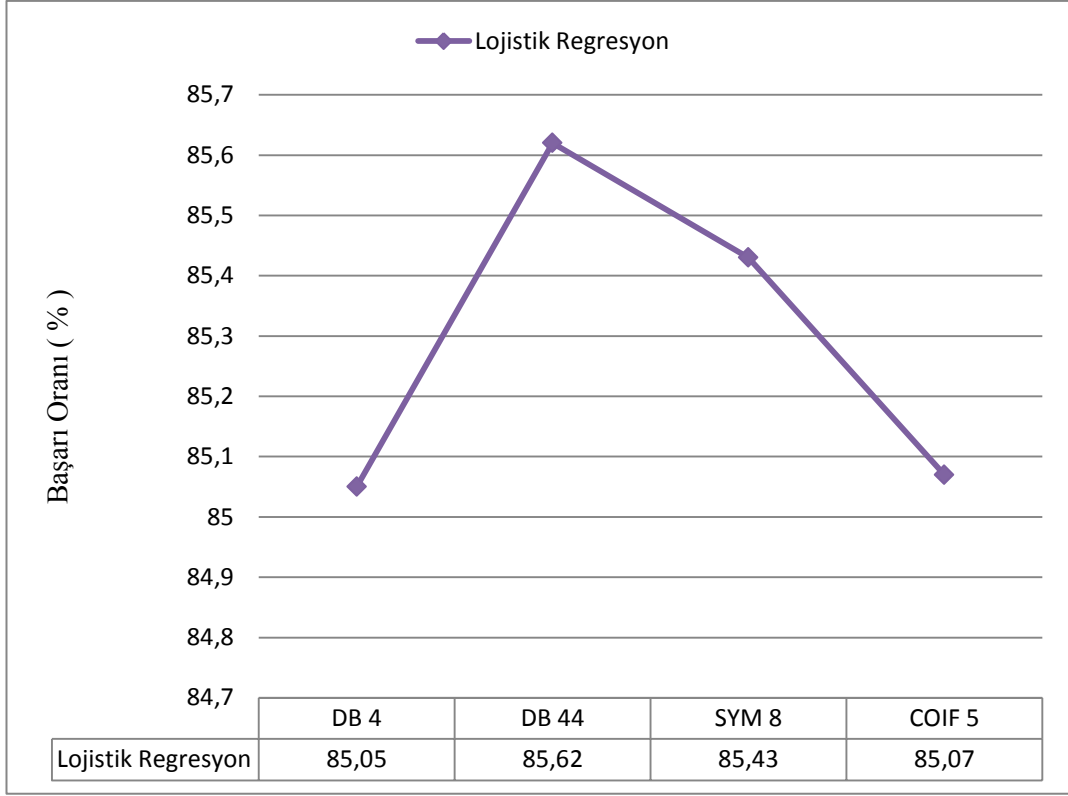
Şekil 3.10'da dalgacık tiplerine göre her bir K değeri için 44 öznitelik arasından sarmal metodun seçtiği öznitelik sayısı değişimi gösterilmektedir. Şekil 3.10'da başarı oranının en yüksek olduğu K=9 değeri Sym 8 dalgacığı için 18 adet, başarı oranının en düşük olduğu K=2 değeri Db 4 dalgacığı için 11 adet öznitelik seçildiği görülmektedir.



Şekil 3.10. Dalgacık tipine göre KNN sınıflandırıcı için seçilen öznelik sayısı

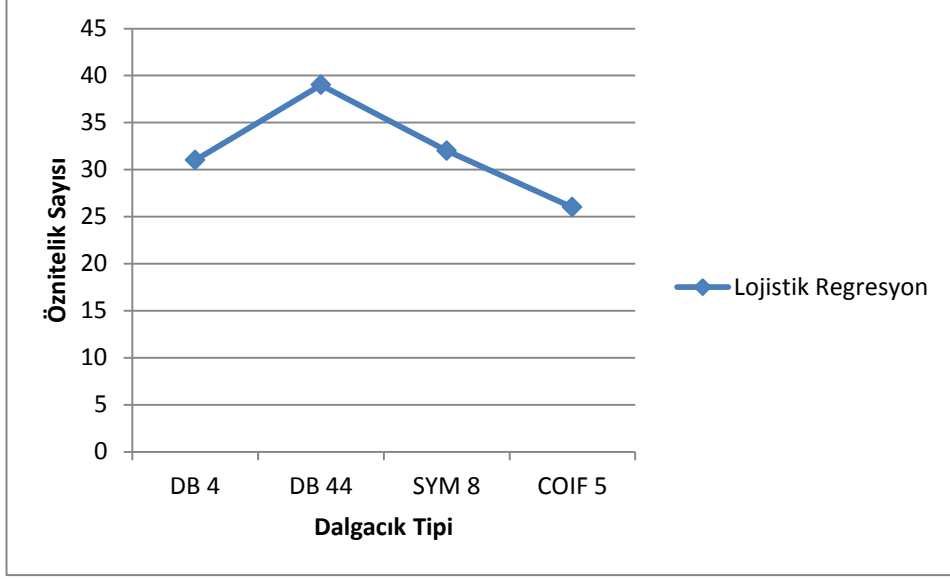
3.5.2. Sarmal metot ile lojistik regresyon

Şekil 3.11’de dalgacık tipine göre lojistik regresyon sınıflandırıcısıyla elde edilen başarı oranı değişimleri gösterilmektedir. Yapılan çalışmada en yüksek başarı oranının Db 44 dalgacıyla % 85,62 olduğu görülmüştür.



Şekil 3.11. Dalgacık tipine göre lojistik regresyon sınıflandırıcısı başarısı

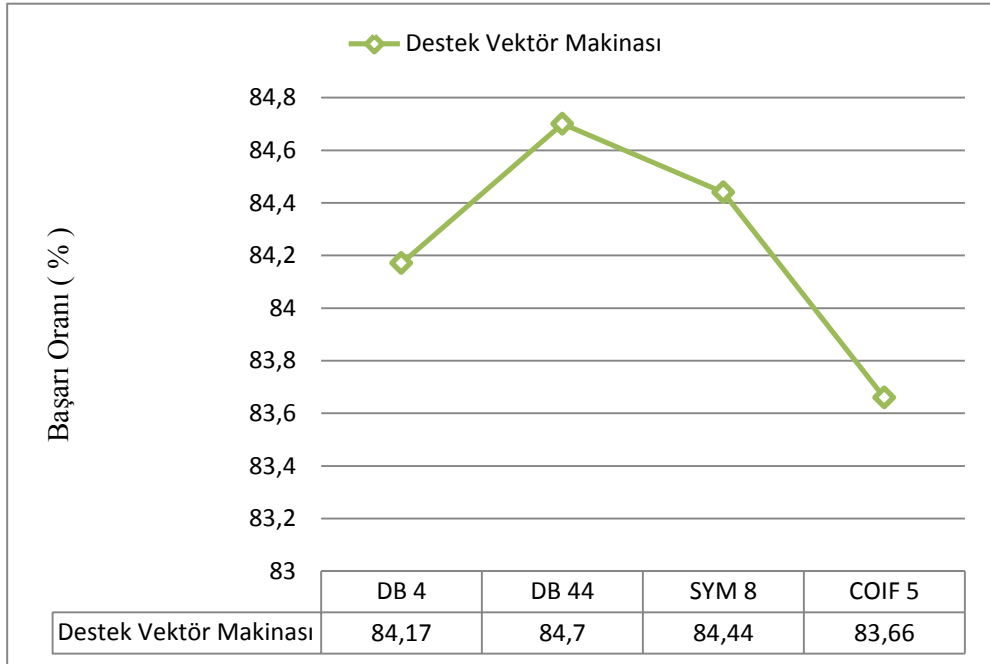
Şekil 3.12’de Lojistik regresyon sınıflandırıcısıyla dalgacık tipine göre sarmal metot ile öznelik azaltımı sonucunda 44 öznelik arasından seçilen özneliklerin sayıları gösterilmiştir. Şekil 3.11 ve 3.12 birlikte değerlendirildiğinde Db 44 dalgacığı için elde edilen %85,62 başarı oranı 39 adet öznelikten, Db 4 dalgacığı için %85,05 başarı oranı 31 adet, Sym 8 dalgacığı için % 85,43 başarı oranı 32 adet, Coif 5 dalgacığı için %85,07 başarı oranı 26 adet öznelikten elde edilmiştir.



Şekil 3.12. Dalgacık tipine göre lojistik regresyon sınıflandırıcısı için seçilen öznitelik sayısı

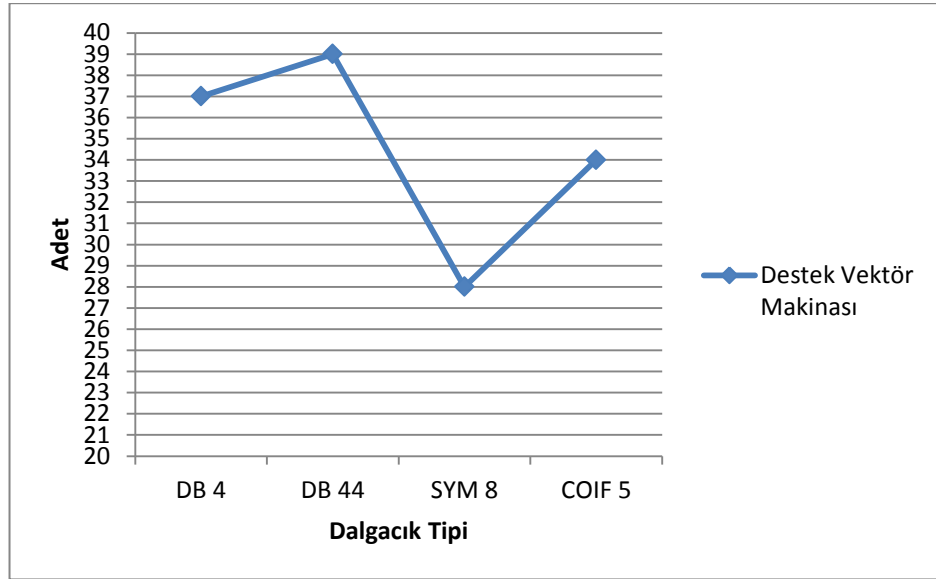
3.5.3. Sarmal metot ile SVM sınıflandırıcısı

Şekil 3.13'de çekirdek fonksiyonu PolyKernel olarak seçilen SVM sınıflandırıcısı, sarmal değerlendirici ve alt küme üretici Bestfirst olarak belirlenmiş, Weka programı aracılığıyla öznitelik azaltımı gerçekleştirilmiştir. Bu çalışma sonucunda en yüksek başarı oranı %84,7 ile Db 44 dalgacığı ile elde edilmiştir.



Şekil 3.13. Dalgacık tipine göre destek vektör makinası sınıflandırıcısı başarıları

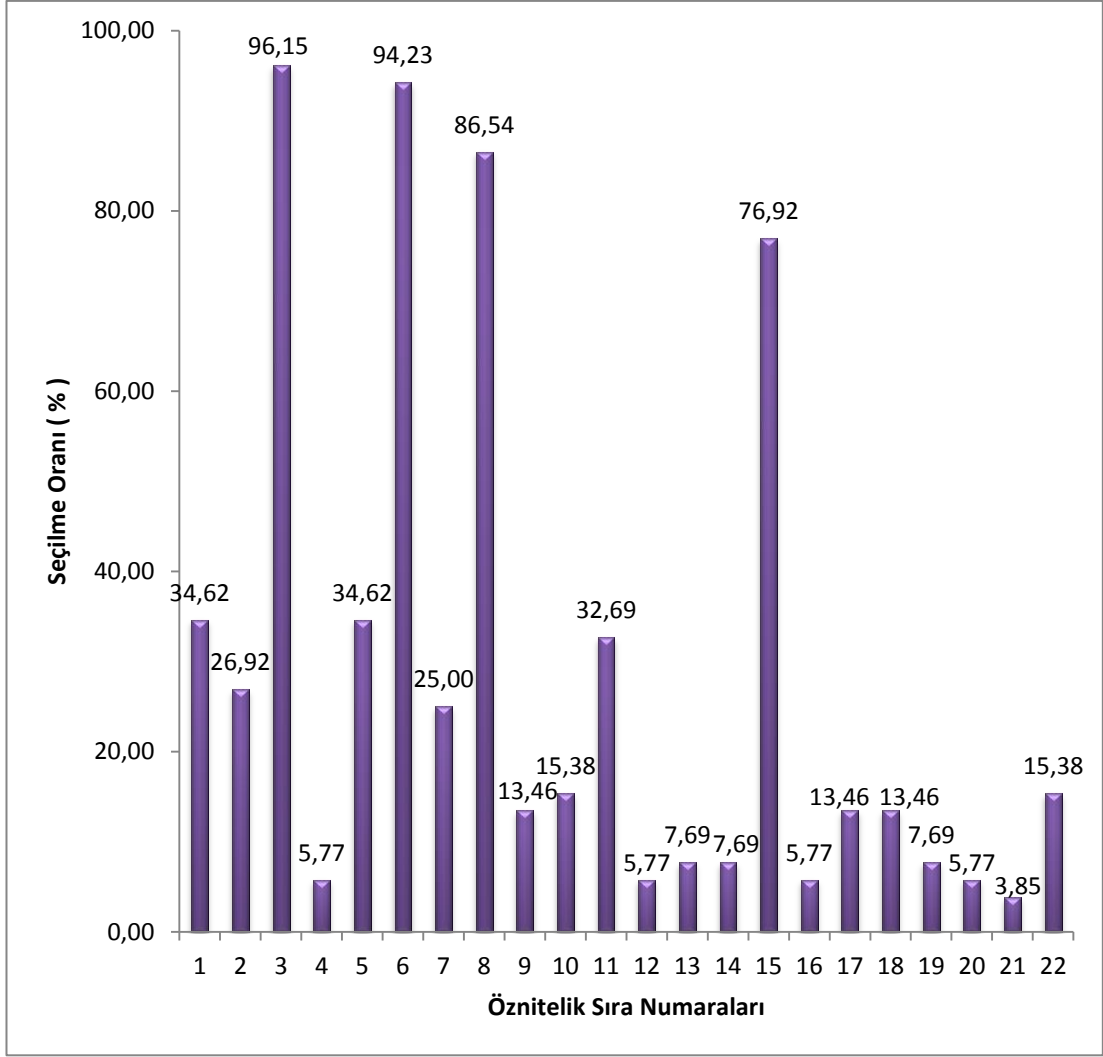
Şekil 3.14’de destek vektör makinaları sınıflandırıcısıyla dalgacık tipine göre sarmal metotla öznelik azaltımı sonucunda 44 öznelik arasından seçilen özneliklerin sayıları gösterilmiştir. Şekil 3.13 ve 3.14 birlikte değerlendirildiğinde Db 44 dalgacığı için elde edilen %84,7 başarı oranı 39 adet öznelikten, Db 4 dalgacığı için %84,17 başarı oranı 37 adet, Sym 8 dalgacığı için % 84,44 başarı oranı 28 adet, Coif 5 dalgacığı için %83,66 başarı oranı 33 adet öznelik ile elde edilmiştir.



Şekil 3.14. Dalgacık tipine göre destek vektör makinası sınıflandırıcı için seçilen öznelik sayısı

3.5.4. Sarmal metot ile ayırt edici öznelikler

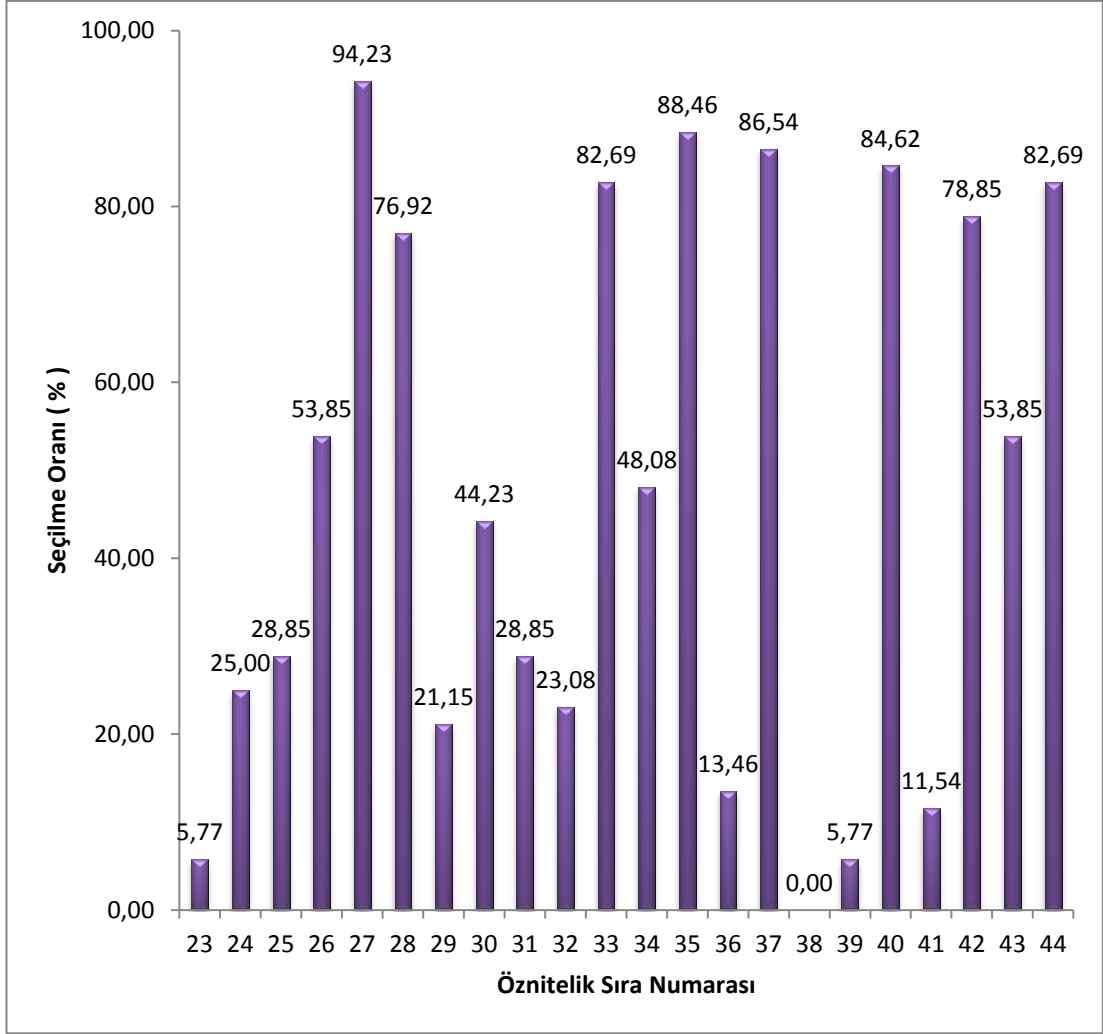
Şekil 3.15’de 44 adet elde edilmiş olan özneliklerin, sarmal metot ile öznelik azaltımı sonucunda KNN $K = 2 \dots 12$, Lojistik regresyon, SVM ile sınıflandırılması sonucunda toplam 13 farklı sınıflandırıcı sonucunda her bir özneliğın seçilme oranı yüzdeler olarak verilmiştir. Şekil 3.15’de verilen sıra numaraları Tablo 3.3 temel alınarak yazılmıştır. İki grafikteki veriler de incelendiğinde uyku evrelerini ayırt eden en belirleyici özneliğın %96,15’lik oranla teta bandı enerjisi / toplam enerji olduğunu, uyku evrelerini ayırt edemeyen özneliğın ise %0’lık oranla 1. Kanal EOG’si olduğu görülmektedir.



Şekil 3.15. Sarmal metoda göre özniteliklerin seçilme oranı

Şekil 3.15’de 44 adet öznitelik arasında sarmal metodun seçtiği en iyi ayırt edici öznitelige sahip öznitelikleri sırasıyla;

Teta Dalgası Enerjisi / Toplam Enerji, epğun minimum genliđi, Delta bandı enerjisi, 5 Hz üstü EOG2 sinyali enerjisi, Beta bandının standart sapması, 1.kanal EOG’sinin standart sapması, 5 Hz üstü EOG1 sinyali enerjisi, Hjorth parametresi 3(complexity), Hjorth parametresi 1(activity), Delta bandının ortalaması, Beta bandı düzensizlik (entropy) değeri olduđu görölmektedir.



Şekil 3.15. (Devam) Sarmal metoda göre özniteliklerin seçilme oranı

3.5.6. Filtre metodu ile KNN sınıflandırıcısı

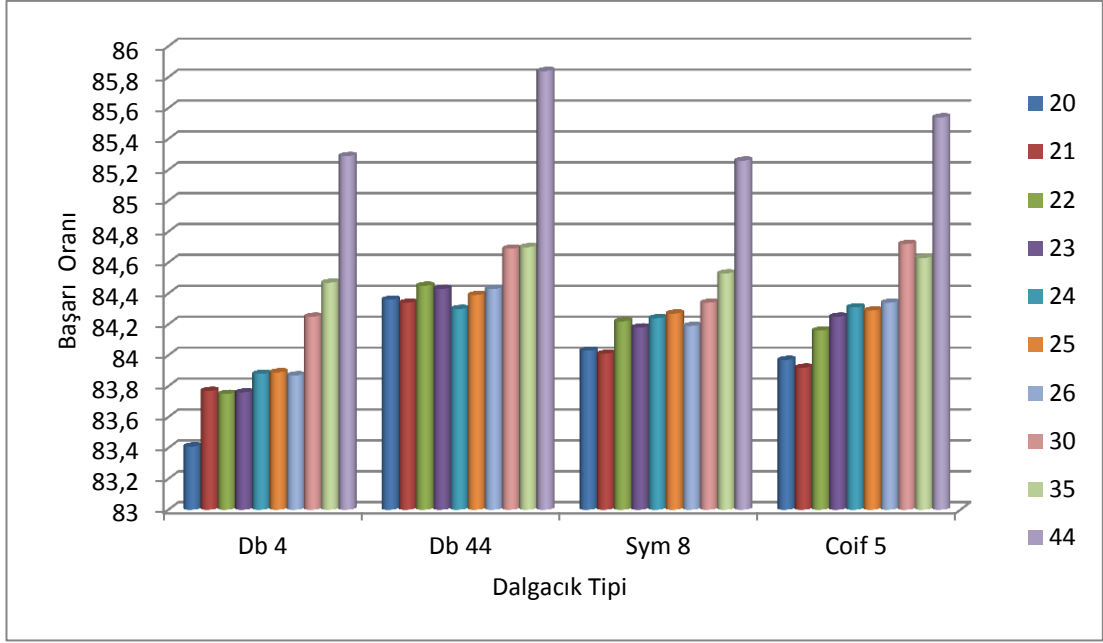
Tablo 3.4’de dalgacık tiplerine göre, seçilen özniteliklerin başarı oranları arasındaki ilişki gösterilmiştir. Yapılan denemeler sonucunda diğerlerine göre daha yüksek başarı oranı sağlayan KNN 5-9-11 sınıflandırıcısı sonuçları Tablo 3.4’de gösterilmiştir. Weka ile yapılan özellik azaltımı için değerlendirici GainRatioAttributeEval, alt küme üretici ranker olarak seçilmiştir.

Tablo 3.4. KNN (5-9-11) sınıflandırıcı öznitelik sayısına göre elde edilen başarı oranları

		Öznitelik Sayısı									
		20	21	22	23	24	25	26	30	35	44
Db 4	KNN 5	80,42	80,6	80,38	80,62	80,63	80,56	80,89	81,38	81,39	82,17
	KNN 9	81,23	81,5	81,24	81,41	81,61	81,61	81,74	82,12	81,91	82,92
	KNN 11	81,3	81,2	81,15	81,21	81,45	81,47	81,59	82,19	82,06	82,85
Db 44	KNN 5	80,72	80,9	81,12	81,07	81,31	81,57	81,55	81,84	81,85	82,54
	KNN 9	81,6	81,6	81,81	82,01	81,64	81,98	81,94	82,71	82,23	83,43
	KNN 11	81,64	81,7	81,94	82,05	81,81	82,07	82,07	82,65	82,35	83,42
Sym 8	KNN 5	80,29	80,3	80,87	81,11	81,61	81,45	81,46	82,31	81,97	82,49
	KNN 9	81,41	81,4	81,65	81,98	81,95	82,14	82,13	82,8	82,6	83,03
	KNN 11	81,53	81,6	81,74	82,84	82,09	82,1	82,09	82,84	82,65	83,06
Coif 5	KNN 5	80,16	80,18	80,66	80,92	81,13	81,18	81,16	81,3	81,61	82,49
	KNN 9	80,73	80,78	81,43	81,81	82,05	82,00	81,99	82,18	82,29	82,95
	KNN 11	80,92	80,95	81,60	81,82	82,18	82,16	82,15	82,6	82,27	83,04

3.5.7. Filtre metodu ile lojistik regresyon sınıflandırıcısı

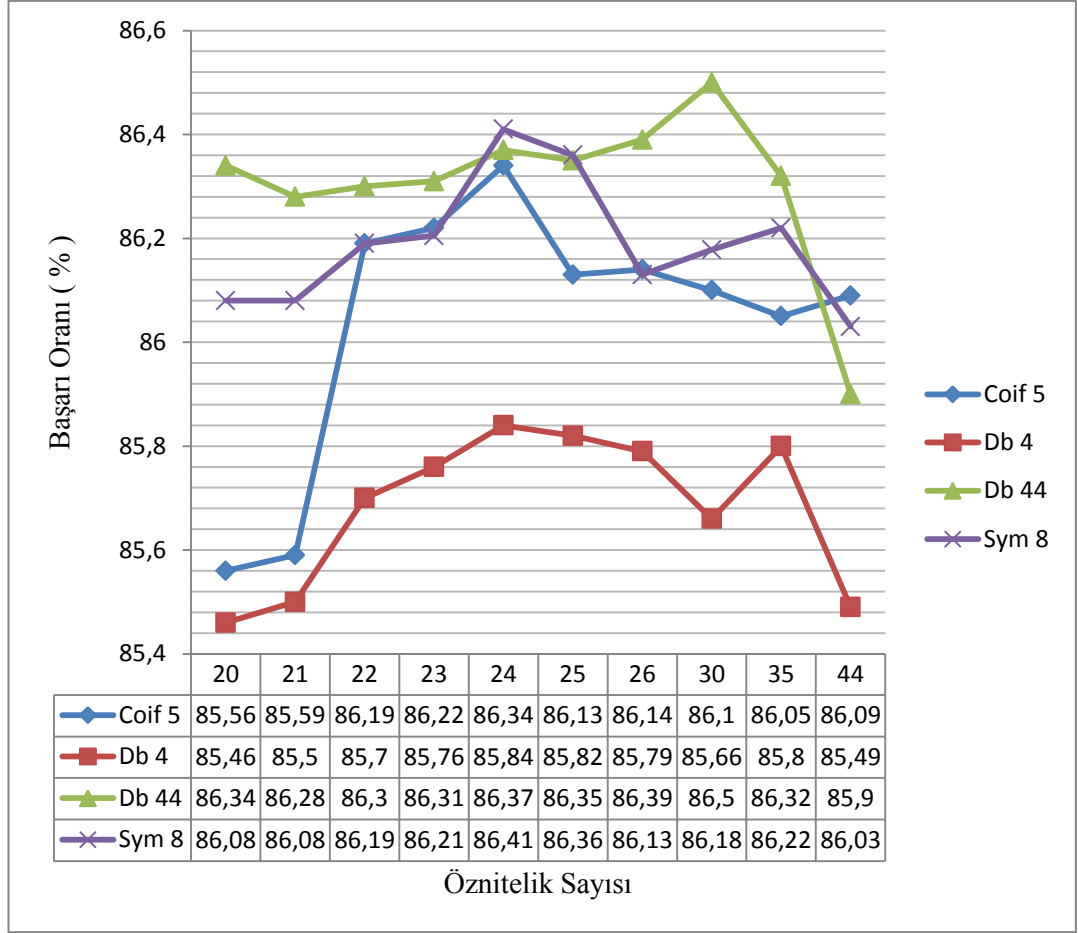
Şekil 3.16'de değerlendirici GainRatioAttributeEval, alt küme üretici Ranker olarak belirlenmiş ve öznitelik azaltımı gerçekleştirilmiştir. Her bir dalgacık için seçilen öznitelik sayısına göre (20, 21, 22, 23, 24, 25, 26, 30, 35, 44) başarı oranları tespit edilmiştir. Filtre metodu ile lojistik regresyon sınıflandırıcısı için öznitelik azaltımının başarıya bir fayda sağlamadığı, öznitelik azaltmadan elde edilen en yüksek başarımın ise Db44 dalgacığı ile %85,8 olduğu Şekil 3.16'dan anlaşılmaktadır.



Şekil 3.16. Filtre metodu lojistik regresyon sınıflandırıcısıyla öznitelik azaltım başarı oranları

3.5.8. Filtre metodu ile SVM sınıflandırıcısı

Şekil 3.17’de çekirdek fonksiyonu PUK olarak seçilen SVM sınıflandırıcısı, değerlendirici SVMAttributeEval ve alt küme üretici Ranker olarak belirlenmiş, Weka programı aracılığıyla öznitelik azaltımı gerçekleştirilmiştir. Ranker’ın belirlediği en yüksek ayırt ediciliğe sahip 20, 21, 22, 23, 24, 25, 26, 30, 35 öznitelik sırasıyla denenmiş ve sonuçları Şekil 3.17’de gösterilmiştir.



Şekil 3.17. Filtre metodu ile SVM sınıflandırıcısı başarı oranı

Şekil 3.17’de Db 44 dalgacığı ile elde edilen özniteliklerin uyku evreleri tespitinde genel olarak daha yüksek bir başarı elde ettiği görülmektedir. 30 öznitelik ile Db 44 dalgacığı %86,5’lik bir başarı oranına ulaşmıştır.

Tablo 3.5’de dalgacık tiplerine göre elde edilen özniteliklerden en yüksek sınıflama başarısı oranına sahip olan öznitelik ile sınıflandırılmış verilerin karmaşıklık matrisi yer almaktadır. Tablo 3.5, Şekil 3.17’deki en yüksek sınıflandırma başarısı dikkate alınarak oluşturulmuştur.

Tablo 3.5. Her bir dalgacık için karmaşıklık matrisi

	a	b	c	d	
Coif 5 Dalgacığı	2314	441	0	15	a = NREM3
	254	5131	449	74	b = NREM2
	2	375	2605	200	c = NREM1+REM
	8	37	130	2507	d = WAKE
Db 4 Dalgacığı	2325	426	0	19	a = NREM3
	275	5109	443	81	b = NREM2
	1	424	2545	212	c = NREM1+REM
	8	42	128	2504	d = WAKE
Db 44 Dalgacığı	2325	429	0	16	a = NREM3
	245	5135	452	76	b = NREM2
	1	373	2603	205	c = NREM1+REM
	6	34	125	2517	d = WAKE
Sym 8 Dalgacığı	2329	427	0	14	a = NREM3
	263	5136	433	76	b = NREM2
	2	388	2581	211	c = NREM1+REM
	7	29	125	2521	d = WAKE

Dalgacık tipine göre karmaşıklık matrisi tablolarda gösterilmiştir.

Tablo 3.6. Coif 5 dalgacığı için karmaşıklık matrisinin türevleri

Doğru Pozitif Oranı	Yanlış Pozitif Oranı	Kesinlik	Recall	F -Measure	Sınıf
0,835	0,022	0,898	0,835	0,865	NREM3
0,868	0,099	0,857	0,868	0,863	NREM2
0,819	0,051	0,818	0,819	0,818	NREM1+REM
0,935	0,024	0,897	0,935	0,915	WAKE
0,863	0,060	0,864	0,863	0,863	Ortalama

Tablo 3.7. Db 4 dalgacıđı için karmaşıklık matrisinin türevleri

Dođru Pozitif Oranı	Yanlıř Pozitif Oranı	Kesinlik	Recall	F -Measure	Sınıf
0,839	0,024	0,891	0,839	0,864	NREM3
0,865	0,103	0,851	0,865	0,858	NREM2
0,800	0,050	0,817	0,800	0,808	NREM1+REM
0,934	0,026	0,889	0,934	0,911	WAKE
0,858	0,062	0,858	0,858	0,803	Ortalama

Tablo 3.8. Db 44 dalgacıđı için karmaşıklık matrisinin türevleri

Dođru Pozitif Oranı	Yanlıř Pozitif Oranı	Kesinlik	Recall	F -Measure	Sınıf
0,839	0,021	0,902	0,839	0,870	NREM3
0,869	0,097	0,860	0,869	0,865	NREM2
0,818	0,051	0,819	0,818	0,818	NREM1+REM
0,938	0,025	0,894	0,938	0,916	WAKE
0,865	0,059	0,865	0,865	0,865	Ortalama

Tablo 3.9. Sym 8 dalgacıđı için karmaşıklık matrisinin türevleri

Dođru Pozitif Oranı	Yanlıř Pozitif Oranı	Kesinlik	Recall	F -Measure	Sınıf
0,841	0,023	0,895	0,841	0,867	NREM3
0,869	0,098	0,859	0,869	0,864	NREM2
0,811	0,049	0,822	0,811	0,817	NREM1+REM
0,940	0,025	0,893	0,940	0,916	WAKE
0,864	0,060	0,864	0,864	0,864	Ortalama



Şekil 3.18. Dalgacık Çeşitlerine göre uyku evresi sınıflandırma başarısı

Şekil 3.18’de 14542 adet epöğü sınıflandırma sonucunda SVM ile yapılan sınıflandırma sonucunda Sym 8 dalgacığı için NREM 3 evresini %84,1, NREM 2 %86,9, NREM1+REM %81,1, WAKE %94 oranında, Db 44 dalgacığı için NREM 3 evresini %83,9, NREM 2 %86,9 NREM1+REM %81,8, WAKE %93,8 oranında başarı sağladığı görülmüştür.

3.5.9. Filtre metoduna göre ayırt edici öznitelikler

Tablo 3.10’da yüksek sınıflama başarı oranına sahip SVM sınıflandırıcısı için Db 44 ve Sym 8 dalgacık katsayıları ile elde edilmiş olan özniteliklerin listesi verilmiştir. 44 tane elde edilmiş olan öznitelikten, alt küme üretici ranker olarak seçilen öznitelik azaltımı ve SVM sınıflandırıcısı sonucunda seçilen öznitelikler verilmiştir.

Uyku evresi tespitinde uyku evrelerini ayırt edicilik bakımından seçilen en başarılı özniteliklere sıra numarasına göre Tablo 3.10’da yer verilmiştir.

Tablo 3.10. Uyku evresi tespitinde başarı oranı yüksek olan öznelikler

Sıra no	Sym 8 Dalgacığı	Db 44 Dalgacığı
	Seçilen öznelik	Seçilen öznelik
1	Delta bandı enerjisi	Delta bandı enerjisi
2	Hjorth parametresi 1(activity)	Beta bandının standart sapması
3	Delta bandının ortalaması	Delta bandının ortalaması
4	Teta bandı enerjisi	Teta bandı enerjisi
5	Epoğun minimum genliği	Epoğun minimum genliği
6	Beta bandının standart sapması	Hjorth parametresi 3(complexity)
7	Teta Dalgası Enerjisi / Toplam Enerji	Teta Dalgası Enerjisi / Toplam Enerji
8	Hjorth parametresi 3(complexity)	Delta bandının standart sapması
9	Delta bandı düzensizlik (entropy) değeri	Delta bandı düzensizlik (entropy) değeri
10	Delta Dalgası Enerjisi / Toplam Enerji	Hjorth parametresi 2(mobility)
11	Hjorth parametresi 2(mobility)	Hjorth parametresi 1(activity)
12	Alfa Dalgası Enerjisi / Toplam Enerji	Alfa Dalgası Enerjisi / Toplam Enerji
13	Alfa bandı enerjisi	Epoğun maksimum genliği
14	Alfa bandının standart sapması	Alfa bandının standart sapması
15	Epoğun maksimum genliği	Alfa bandı enerjisi
16	Delta bandının standart sapması	Teta bandının standart sapması
17	Teta bandı düzensizlik (entropy) değeri	Delta bandının maksimum genlik değeri
18	Teta bandının standart sapması	Teta bandı düzensizlik (entropy) değeri
19	2.kanal EOG' sinin standart sapması	Delta Dalgası Enerjisi / Toplam Enerji
20	Beta bandının minimum genlik değeri	Beta bandının maksimum genlik değeri
21	5 Hz üstü EOG 1 sinyali Enerjisi	2.kanal EOG' sinin standart sapması
22	Alfa bandı düzensizlik (entropy) değeri	5 Hz üstü EOG 1 sinyali Enerjisi
23	Delta bandının maksimum genlik değeri	Epoğun Standart Sapması
24	Epoğun Standart Sapması	Alfa bandı düzensizlik (entropy) değeri
25	-	Alfa bandının minimum genlik değeri
26	-	Beta Dalgası Enerjisi / Toplam Enerji
27	-	5 Hz üstü EOG 2 sinyali Enerjisi
28	-	Alfa bandının maksimum genlik değeri
29	-	5 Hz altı EOG 2 sinyali Enerjisi
30	-	1.kanal EOG' sinin tepeden tepeye genliği

4. SONUÇLAR ve ÖNERİLER

Bu çalışmada, AASM standartları ile belirlenmiş olan uyku evreleri tespitinin bilgisayar ortamında yapılması amaçlanmıştır. Veriler dreamsproject veritabanından alınmıştır. Bu veriler 18 farklı hastadan Belçika hastanesi uyku laboratuvarından 32 kanallı PSG cihazı ile en az iki EOG kanalı, 3 EEG kanalı ve bir tane EMG kanalından alınmıştır. Veriler edf formatında kaydedilmiştir. NREM 1 ve REM evrelerinin ayırt edici özellikleri birbirlerine çok benzediğinden dolayı bu iki evre tek evre olarak göz önünde bulundurulmuş ve sınıflandırma yapılmıştır.

Uyku evreleri AASM kriterlerine göre ise 30 sn'lik epoklar halinde skorlanmıştır. Skorlamalar uyku laboratuvarında uzman tarafından yapılmıştır. Bu veriler dreamsproject veritabanında metin dosyası olarak yer almaktadır. Elde edilen bu veriler uyku evreleri skorlamasında referans alınmıştır ve doğruluk oranları bu verilere dayanarak tespit edilmiştir.

Her bir hasta için ortalama 8 saat süren uyku süresi 30 sn'lik epoklara bölünmüş, farklı dalgacık çeşitleri ile dalgacık analizi yapılarak alt frekans bantları elde edilmiş ve her bir epok için 44 adet öznitelik elde edilmiştir.

Uyku evresi sınıflandırmasında daha başarılı bir sonuç elde etmek için çıkartılan öznitelikler farklı metotlarla azaltılarak doğruluk oranının artırılması amaçlanmıştır. Bu amaç gerçekleştirilirken de özellik azaltım yöntemleri arasındaki farklılıklar ve başarı oranları karşılaştırılmış aynı zamanda derinlemesine incelenmiştir.

Coif 5 ve Db 4 dalgacığı ortalama başarı oranı %85 gibi yüksek bir değer olmasına rağmen Sym 8 ve Db 44 dalgacıklarına göre başarı oranları daha düşüktür.

Sym 8 ve Db 44 dalgacıkları kullanılarak elde edilen özniteliklerin azaltımı sonucunda, öznitelik azaltımı olmadan uyku evreleri sınıflandırma başarı oranı Db 44 dalgacığı için %85,90'dan %86,5, Sym 8 dalgacığı için %86,03'den %86,40 başarı oranına ulaşılmış aynı zamanda kullanılan öznitelik sayısının da azaltılmasıyla işlem yükü hafifletilmiştir.

WAKE evresi tespitinde elde edilen başarı oranı Db 44 dalgacığı için %93,8, Sym 8 dalgacığı için ise % 94 olarak tespit edilmiştir. En düşük sınıflandırma başarı oranına sahip olan NREM1+REM evresi başarı oranı ise Db 44 dalgacığı için %81,8, Sym 8 dalgacığı için ise %81,1'dir. NREM 1 evresi uyanıklık ile uyku arasındaki evre olmasından dolayı NREM1 evresine ait epokların WAKE ya da NREM2 evresi olarak skorlanmasından dolayı başarı oranının diğerlerine göre düştüğü gözlemlenmiştir.

İlerleyen çalışmalarda NREM1 ve REM evrelerinin ayrı ayrı tespiti için sürekli dalgacık yöntemi kullanılarak, bu evrelere özgü bazı özel dalga formlarının tespiti için bir çalışma yapılabilir. Elde edilen özniteliklerin azaltımında sarmal metodun kullanılması çok uzun zaman aldığından daha yüksek performans veren bir cihazla tekrar çalıştırılabilir ve daha kısa sürede sonuçlar elde edilerek farklı çekirdek fonksiyonları denenip sınıflandırma doğrulukları arasında karşılaştırmalar yapılabilir.

KAYNAKLAR

- [1] Pıçak R., İsmailoğulları S., Mazıcıoğlu M. M., Üstünbaş H. B., Aksu M. Birinci Basamakta Uyku Bozukluklarına Yaklaşım ve Öneriler, *Tjfmpe*, 2010, 4(3), 12-22.
- [2] Şahin L., Aşçıoğlu M., Sleep and Regulation of Sleep, *Journal of Health Sciences*, 2013, 22(1), 93-98.
- [3] Uçar M., Bozkurt M., Polat K., Bilgin C., Effect of Digital Filtering to Sleep Stage Classification Using EEG Signals, *Elektronik – Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu*, Bursa, Türkiye, 27-29 Kasım 2014.
- [4] K. Susmakova, Human Sleep and Sleep EEG, *Measurement Science Review*, 2004, 4(2), 59-74.
- [5] Çiftçi T. U., Obstrüktif Uyku Apne Sendromu Tanı ve Tedavi Uzlaş Raporu, *Türk Toraks Dergisi*, UCID 1302-7808, 18-20, 2012.
- [6] Güven S., Scoring of the Respiratory Recordings, *Türkiye Solunum Araştırmaları Derneği*, DOI: 10.5152/Solunum.2013.032.
- [7] Öçal Ö., Acıbadem Maslak Hastanesi Beslenme ve Diyet Polikliniğine Başvuran Yetişkin Bireylerde Besin Tüketiminin Pittsburgh Uyku Kalitesi Ölçeği İle İlişkisi, Yüksek Lisans Tezi, Başkent Üniversitesi, Sağlık Bilimleri Enstitüsü, 2015, 383128.
- [8] Aşık M., Bostancı A., Turhan M., Comparison of Manual and Automated Scoring Techniques in Polysomnography, *Turkish Archives of Otolaryngology*, DOI:10.5152/Tao.2014.422.
- [9] Demir A., Keyf F., Başoğlu Ö., Obstrüktif Uyku Apne Sendromu Tanı ve Tedavi Uzlaş Raporu, *Türk Toraks Dergisi*, UCID 1302-7808, 21-26, 2012.
- [10] The AASM 09657220-4, Manuel for the Scoring of Sleep, *American Academy of Sleep Medicine*, Westchester, 2007.
- [11] Bolat E. D., Tuncer E., EEG Signal Based Sleep Stage Classification Using Discrete Wavelet Transform, *International Conference On Chemisrty, Biomedical and Environment Engineering*, Antalya, Türkiye, 08-11 Ekim 2014.
- [12] Pourzare S., EEG İşaretlerindeki Göz ve Çene Hareketi Artifaktlarının Sınıflandırılması, Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Trabzon, 2012, 321469.

- [13] Köktürk O., Scoring of Sleep Recordings, Türkiye Solunum Araştırmaları Derneği, *Solunum 2013*, DOI: 10.5152/Solunum.2013.031.
- [14] <http://uykubozuklugu.uludag.edu.tr/dersnotlari001.htm>, (Ziyaret Tarihi: 05 Mayıs 2015).
- [15] <http://www.mathworks.com/help/wavelet/gs/continuous-wavelettransform.html?searchHighlight=wavelet%20shifting>, (Ziyaret Tarihi: 07 Mayıs 2015).
- [16] Duman F., EEG Sinyallerinin Analizi ile Uyku Durumunun Belirlenmesi, Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2005, 170016.
- [17] <http://wavelets.pybytes.com/>, (Ziyaret Tarihi: 07 Mayıs 2015).
- [18] Abbak R., Jeodezide Zaman Dizilerinin Dalgacık (Wavelet) Analizi, Doktora Semineri, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya, 2007.
- [19] <http://www.mathworks.com/help/wavelet/gs/continuous-wavelet-transform.html#bsorann-1>, (Ziyaret Tarihi: 07 Mayıs 2015).
- [20] Kabataş M. B., Suspended Sediment Load Prediction by Using Wavelet-Fuzzy Logic Combination Model, M.Sc. Thesis, Istanbul Technical University, Fen Bilimleri Enstitüsü, İstanbul, 2014, 356144.
- [21] Üstündağ M., ECG Sinyallerinde Gürültü Gidermek için Dalgacık Dönüşümünün Fpga Tabanlı Donanımsal Gerçeklemesi, *Fırat Üniv. Mühendislik Bilimleri Dergisi*, 2013, **25**(2), 63-68.
- [22] Açoğlu E. A., Sayısal Süzgeç Tasarımı ve Uygulamaları, Yüksek Lisans Tezi, Başkent Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2008, 216705.
- [23] Yıldırım B., The Effect of Turkish Mortality Improvements on the Cost of Annuities Using Entropy Measure, Yüksek Lisans, Hacettepe University, Institute of Graduate School of Science and Engineering, Ankara, 2014, 372824.
- [24] Lassner C., Lienhart R., Norm-Induced Entropies For Decision Forests, 2015 *IEEE Winter Conference on Applications of Computer Vision*, DOI 10.1109/WACV.2015.134.
- [25] Avşar E., Tek-Sınıf Destek Vektör Makineleri Kullanılarak Epileptik EEG İşaretlerinin Sınıflandırılması, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2009, 251692.
- [26] Vidaurre C., Kramer N., Blankertz B., Schlogl A., Time Domain Parameters as a Feature for EEG-Based Brain Computer Interfaces, *Elsevier Neural Networks*, DOI:10.1016/J.Neunet.2009.07.020.

- [27] Liu X., Shang L., A Fast Wrapper Feature Subset Selection Method Based on Binary Particle Swarm Optimization, *2013 IEEE Congress on Evolutionary Computation*, DOI: 978-1-4799-0454-9/13.
- [28] Nnamoko N. A., Arshad F., England D., Vora J., Norman J., Evaluation of Filter and Wrapper Methods for Feature Selection in Supervised Machine Learning, *PGNet*, 2014, **27**(4), 74-79.
- [29] Kaya M., Gen İfade Verilerinde Öznitelik Seçimi ve Sınıflandırma, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2014, 352621.
- [30] http://www.slideshare.net/ozgur_dolgun/destek-vektör-makineleri, (Ziyaret Tarihi: 01 Mayıs 2015).
- [31] Albayrak A. S., Yılmaz S. K., Karar Ağacı Algoritmaları ve İmkb Verileri Üzerine Bir Uygulama, *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 2009, **14**(1), 31-52.
- [32] Özçakır F. C., Çamurcu A. Y., Birlikte Kuralı Yöntemi için Bir Veri Madenciliği Yazılımı Tasarımı ve Uygulaması, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 2007, **12**(6), 21-37.
- [33] Ayık Y. Z., Özdemir A., Yavuz U., Lise Türü ve Lise Mezuniyet Başarısının Kazanılan Fakülte ile İlişkinin Veri Madenciliği Tekniği ile Analizi, *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2007, **33**(4), 441-454.
- [34] Murugappan M., Human Emotion Classification Using Wavelet Transform and Knn, *International Conference on Pattern Analysis and Intelligent Robotics*, Putrajaya, Malaysia, 28-29 Haziran 2011.
- [35] Isa R. M., Pasya I., Taib M. N., Jahidin A. H., Omar W. R. W., Fuad N., Norhazman H., EEG Brainwave Behaviour Due to Rf Exposure Using Knn Classification, *IEEE 3rd International Conference on System Engineering and Technology*, Shah Alam, Malaysia, 19-20 Augustos 2013.
- [36] İlhan İ., Genom Çaplı İlişki Çalışmaları İçin Yapay Zekâ Teknikleri İle Etiket Snp Seçimi, Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya, 2013, 327021.
- [37] Akşehirli Ö., Tıbbi Araştırmalarda Destek Vektör Makinelerinin Kullanımı, Yüksek Lisans Tezi, Düzce Üniversitesi, Sağlık Bilimleri Enstitüsü, Düzce, 2012, 340813.
- [38] Topaloğlu M., Çevrimiçi Destek Vektör Makineleri Tabanlı Model Öngörülü Denetim, Yüksek Lisans Tezi, Pamukkale Üniversitesi, Fen Bilimleri Enstitüsü, Denizli, 2014, 376204.
- [39] Sahak R., Mansor W., Khuan L. Y., Yassin A. I. M., Zabidi A., Yasmin F., Rahman A., Choice for a Support Vector Machine Kernel Function for

Recognizing Asphyxia from Infant Cries, *IEEE Symposium on Industrial Electronics and Applications*, Kuala Lumpur, Malaysia, 4-6 Ekim 2009.

- [40] Abkar A. K., Yu C., Performance of Svm Based on Puk Kernel in Comprasion to Svm Based on Rbf Kernel in Prediction of Yarn Tenacity, *Indian Journal of Fibre &Textile Research*, 2014, **39**(4), 55-59.
- [41] Üstün B., Melssen W. J., Buydens L. M. C., Facilitating The Application of Support Vector Regression by Using a Universal Pearson VII Function Based Kernel, *Chemometrics and Intelligent Laboratory Systems* Doi:10.1016/J.Chemolab.2005.09.003.
- [42] Budak İ., Veri Madenciliği Teknikleri ile Bilgisayar Ağlarında Anomali Tespiti, Yüksek Lisans Tezi, Karabük Üniversitesi, Fen Bilimleri Enstitüsü, Karabük, 2013, 332402.
- [43] Karakuş K., Veri Madenciliği Teknikleri ile Mobil Telekom Sektöründe Müşterilerin Kredi Skorlamasına İlişkin İstatistiksel Bir Analiz, Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, 2009, 240365.
- [44] Xi R., Lin N., Chen Y., Compression and Aggregation for Logistic Regression Analysis in Data Cubes, *IEEE Transactions on Knowledge and Data Engineering*, 2009, **21**(4), 479-492.
- [45] Turna F., Veri Madenciliği Teknikleriyle Tramvay Arıza, Kayıtlarından Kural Çıkarımı, Yüksek Lisans Tezi, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri, 2011, 300032.
- [46] Gula L., Information Data Management for the Future of Communication, Master's Thesis, Lulea University, Engineering Technology Computer Science and Engineering, Sweden, 2009.
- [47] <http://www.tcts.fpms.ac.be/~devuyst/Databases/DatabaseSubjects/>, (Ziyaret Tarihi: 01 Mayıs 2015).

KİŞİSEL YAYIN ve ESERLER

- [1] Bolat E. D., **Tuncer E.**, EEG Signal Based Sleep Stage Classification Using Discrete Wavelet Transform, *International Conference on Chemistry, Biomedical and Environment Engineering*, Antalya, Türkiye, 08-11 Ekim 2014.

ÖZGEÇMİŞ

1985 yılında Kocaeli’de doğdu. İlk, orta ve lise öğrenimini Kocaeli’de tamamladı. 2002 yılında girdiği Gazi Üniversitesi Teknik Eğitim Fakültesi Elektronik Öğretmenliği Bölümü’nden 2008 yılında Elektronik Öğretmeni olarak mezun oldu. 2010 yılından beri Milli Eğitim Bakanlığında Elektronik Öğretmeni olarak görev yapmaktadır.