

**KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ
ANABİLİM DALI**

YÜKSEK LİSANS TEZİ

**EEG İŞARETLERİNDE EPİLEPSİ NÖBET TAHMİNİ VE
TESPİTİ**

ALİYA ZHUNİS

KOCAELİ 2019

KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

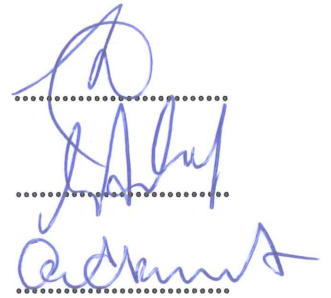
ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ
ANABİLİM DALI

YÜKSEK LİSANS TEZİ

EEG İŞARETLERİNDE EPİLEPSİ NÖBET TAHMİNİ VE
TESPİTİ

ALİYA ZHUNİS

Dr.Öğr.Üyesi Aysun TAŞYAPI ÇELEBİ
Danışman, Kocaeli Üniversitesi
Dr.Öğr.Üyesi Gökçen ÇETİNEL
Jüri Üyesi, Sakarya Üniversitesi
Dr.Öğr.Üyesi Orhan AKBULUT
Jüri Üyesi, Kocaeli Üniversitesi



Handwritten signatures of the jury members, including the advisor and two jury members, positioned to the right of the text.

Tezin Savunulduğu Tarih: 18.07.2019

ÖNSÖZ VE TEŞEKKÜR

Teknolojilerin gelişmesi sayesinde epilepsi hastalıklarının teşhis edilmesinde çeşitli çalışmalar yapılmaktadır.

Bu tez çalışmasında, epilepsi hastaları için EEG sinyalleri kullanılarak Ayrık Dalgacık Dönüşümü ve Ampirik Kip Ayırışımı temelli epilepsi nöbet tahmini ve tespiti yapılmıştır. Bu değerlendirme, epilepsi hastalığının teşhisi konusunda pratik bir epilepsi nöbet tahmin ve tespit sisteminin tasarlanmasında kullanılabilir.

Bu çalışmayı yapmama yardım eden, çalışmanın her aşamasında ilgi ve desteğini esirgemeyen değerli hocam Dr. Öğr. Üyesi Aysun Taşyapı ÇELEBİ'ye, arkadaşlarıma ve bana hem maddi hem manevi katkıda destekleyen aileme teşekkürlerimi sunarım.

Haziran – 2019

Aliya ZHUNİS

İÇİNDEKİLER

ÖNSÖZ VE TEŞEKKÜR	i
İÇİNDEKİLER	ii
ŞEKİLLER DİZİNİ	iii
TABLolar DİZİNİ	v
SİMGELER VE KISALTMALAR DİZİNİ	vi
ÖZET	vii
ABSTRACT	viii
GİRİŞ	1
1. ELEKTROENSEFALOGRAFİ (EEG)	2
1.1. Elektroensefalogram İşaretlerinin Ölçüm Yöntemi	4
1.2. EEG'nin Kullanım Alanları	5
1.3. Epilepsi	5
2. LİTERATÜR TARAMASI	6
3. MATERYAL VE YÖNTEM	10
3.1. Materyal	10
3.2. Özellik Çıkartımı	11
3.2.1. Ayrık dalgacık dönüşümü (ADD)	12
3.2.2. Ampirik kip ayrışımı (AKA)	19
3.2.3. İstatistiksel momentler	22
3.3. Sınıflandırma	24
3.3.1. Destek vektör makineleri (DVM)	24
3.3.1.1. Lineer ayrılma durumu	26
3.3.1.2. Lineer ayrılamama durumu	28
3.3.1.3. Lineer olmayan DVM	29
3.3.2. K-en yakın komşuluk (K-YK)	29
3.3.3. Karar Ağacı (KA)	30
4. DENEYSEL SONUÇLAR	33
4.1. EEG Verilerinden Özellik Çıkartımı	33
4.2. EEG İşaretlerinde Epilepsi EEG Tespiti ve Tahmini	36
4.3. EEG İşaretlerine Dair Analiz Sonuçları	37
5. SONUÇLAR VE ÖNERİLER	61
KAYNAKLAR	62
KİŞİSEL YAYIN VE ESERLER	65
ÖZGEÇMİŞ	66

ŞEKİLLER DİZİNİ

Şekil 1.1.	Elektrot yerleştirme sistemi.....	4
Şekil 3.1.	EEG sinyalleri: A- sağlıklı; D- nöbet öncesi; E- nöbet anı verileri.....	10
Şekil 3.2.	Epilepsi nöbet tahmini ve tespiti için blok şeması	11
Şekil 3.3.	Zamana bağlı ana DD'nün skalandırılması ile elde edilen temel fonksiyonlarının spektrumları (2 çarpanı ile).....	14
Şekil 3.4.	Sinyal spektrumunun sonlu sayıda spektrum ile kaplanması sağlanması	15
Şekil 3.5.	İteratif filtre kümesi ile sinyal spektrumunun frekans bantlarına ayrılması.....	16
Şekil 3.6.	İteratif filtreleme işleminin gerçekleşmesi	19
Şekil 3.7.	Ampirik Kip Ayrışımına ait bir görsel.....	21
Şekil 3.8.	Dağılımlar ve asimetri katsayıları	23
Şekil 3.9.	Dağılımlar ve basıklık katsayıları.....	23
Şekil 3.10.	Optimum ayırıcı düzlemler	25
Şekil 3.11.	Makine öğrenmesi sınıflandırma mimarisi	25
Şekil 3.12.	DVM sınıflandırıcısı.....	26
Şekil 3.13.	Lineer ayrılabilir durumda optimum ayırıcı aşırıdüzlem.....	28
Şekil 3.14.	Giriş uzayının çok boyutlu uzaya dönüşümü	29
Şekil 3.15.	K-YK adımları.....	30
Şekil 3.16.	Karar Ağacı örneği	31
Şekil 4.1.	Nöbet anındaki EEG sinyalinden ADD yöntemiyle elde edilen alt bantlar	34
Şekil 4.2.	A kümesine ait sağlıklı EEG döneminin ilk beş İKF'i	34
Şekil 4.3.	D kümesine ait nöbet öncesi EEG döneminin ilk beş İKF'i	35
Şekil 4.4.	E kümesine ait nöbet anı EEG döneminin ilk beş İKF'i.....	35
Şekil 4.5.	Gözlenen örnek sayısına göre durum I için KA modeline ait karışım matrisi	39
Şekil 4.6.	Durum I için KA modeline ait karışım matrisi	40
Şekil 4.7.	Nöbet anı pozitif sınıfı olduğunda durum I için KA sınıflandırıcısına ait ROC ve AUC eğrileri.....	41
Şekil 4.8.	Gözlenen örnek sayısına göre DVM sınıflandırıcısına ait II. durum için karışım matrisi	42
Şekil 4.9.	Durum I için DVM sınıflandırıcısına ait karışım matrisi.....	43
Şekil 4.10.	Nöbet anı pozitif sınıfı olduğunda durum II için DVM modeline ait ROC ve AUC eğrileri.....	43
Şekil 4.11.	Gözlenen örnek sayısına göre durum III için DVM sınıflandırıcısına ait karışım matrisi	44
Şekil 4.12.	Durum III için DVM sınıflandırıcısına ait karışım matrisi.....	45
Şekil 4.13.	Nöbet anı pozitif sınıfı olduğunda durum III için DVM sınıflandırıcısına ait ROC ve AUC eğrileri.....	46
Şekil 4.14.	Gözlenen örnek sayısına göre durum IV, DVM sınıflandırıcı, karışım matrisi	46

Şekil 4.15. Durum IV, DVM sınıflandırıcı, karışım matrisi.....	47
Şekil 4.16. Nöbet anı pozitif sınıfı olduğunda DVM sınıflandırıcı, durum IV için ROC ve AUC eğrileri.....	48
Şekil 4.17. Gözlenen örnek sayısına göre KA, durum V için karışım matrisi	49
Şekil 4.18. KA, durum V için karışım matrisi	49
Şekil 4.19. Nöbet anı pozitif sınıfı olduğunda V. durum için KA modeline ait ROC ve AUC eğrileri	50
Şekil 4.20. Gözlenen örnek sayısına göre durum I, DVM modeli, karışım matrisi	51
Şekil 4.21. Durum I, DVM modeli, karışım matrisi	52
Şekil 4.22. Nöbet anı pozitif sınıfı olduğunda durum I için DVM modeli, ROC ve AUC eğrileri	52
Şekil 4.23. Gözlenen örnek sayısına göre durum II, KA modeli, karışım matrisi	53
Şekil 4.24. Durum II, KA modeli, karışım matrisi	54
Şekil 4.25. Nöbet anı pozitif sınıfı olduğunda durum II için KA modeli, ROC ve AUC eğrileri	55
Şekil 4.26. Gözlenen örnek sayısına göre III. durum için KA modeli, karışım matrisi	56
Şekil 4.27. Durum III, KA modeli, karışım matrisi	56
Şekil 4.28. Nöbet anı pozitif sınıfı olduğunda durum III için KA modeli, ROC ve AUC eğrileri	57
Şekil 4.29. Gözlenen örnek sayısına göre durum IV, KA modeli, karışım matrisi	57
Şekil 4.30. Durum IV, KA modeli, karışım matrisi	58
Şekil 4.31. Nöbet anı pozitif sınıfı olduğunda durum IV, KA modeli, ROC ve AUC eğrileri.....	58
Şekil 4.32. Gözlenen örnek sayısına göre durum V, KA modeli, karışım matrisi	59
Şekil 4.33. Durum V, KA modeli, karışım matrisi	60
Şekil 4.34. Nöbet anı pozitif sınıfı olduğunda durum V, KA modeli, ROC ve AUC eğrileri.....	60

TABLULAR DİZİNİ

Tablo 1.1. EEG işaretlerinin kapsadıkları bantlar	3
Tablo 4.1. EEG sinyallerinin 4 dereceli ADD ile ayrıştırıldığında elde edilen detay (D) ve yaklaşık (A) katsayıları.....	33
Tablo 4.2. İKF'lerin kapsadığı frekans aralığı.....	33
Tablo 4.3. AKA yöntemi ile hesaplanan E kümesine ait istatistiksel momenter	36
Tablo 4.4. Farklı senaryoların açıklanması.....	36
Tablo 4.5. ADD yöntemine göre elde edilen doğruluk değerleri.....	36
Tablo 4.6. AKA yöntemine göre elde edilen doğruluk değerleri.....	37
Tablo 4.7. ADD yöntemine göre tüm bantlar için elde edilen doğruluk değerleri	37
Tablo 4.8. AKA yöntemine göre tüm İKF'ler için elde edilen doğruluk değerleri	37
Tablo 4.9. Karışım matrisi	37

SİMGELER VE KISALTMALAR DİZİNİ

Kısaltmalar

ADD	: Ayrık Dalgacık Dönüşümü
AKA	: Ampirik Kip Ayrışımı
AUC	: Area Under Curve (Eğri Altındaki Alan)
BDE	: Bağlı Dalga Enerjisi (Relative Wave Energy)
DD	: Dalgacık Dönüşüm
DESA	: Derin Evrişimli Sinir Ağları (Deep Convolutional Neural Network)
DSA	: Derin Sinir Ağları (Deep Neural Network)
DVM	: Destek Vektör Makineleri
DWT	: Discrete Wavelet Transform (Ayrık Dalgacık Dönüşümü)
EEG	: Electroencephalography (Elektroansefalografi)
EMD	: Empirical Mode Decomposition (Ampirik Kip Ayrışımı)
EMG	: Electromyography (Elektromiyografi)
FN	: False Negative (Yanlış Negatif)
FP	: False Positive (Yanlış Pozitif)
IMF	: Intrinsic Mode Functions (İçkin Kip Fonksiyonu)
İKF	: İçkin Kip Fonksiyonu
KA	: Karar Ağacı
KDA	: Kuadratik Diskriminant Analizi (Quadratic Discriminant Analysis)
K-NN	: K-Nearest Neighbor (K-en Yakın Komşuluk)
K-YK	: K-en Yakın Komşuluk
LDA	: Lineer Diskriminant Analizi (Linear Discriminant Analysis)
MATLAB	: Matrix Laboratory (Matris Laboratuvarı)
NBS	: Naive Bayes Sınıflandırıcısı (Naive Bayes Classifier)
ROC	: Receiver Operating Characteristic (Alıcı İşletim Karakteristiği)
SVM	: Support Vector Machine (Destek Vektör Makineleri)
TN	: True Negative (Gerçek Negatif)
TP	: True Positive (Gerçek Pozitif)
YSA	: Yapay Sinir Ağı

EEG İŞARETLERİNDE EPİLEPSİ NÖBET TAHMİNİ VE TESPİTİ

ÖZET

Epilepsi nöbeti beynin sinir hücrelerinde anormal elektriksel boşalım olarak meydana gelen nörolojik bir bozukluktur. Elektroansefalografi (EEG) ise bu anormal elektriksel aktiviteyi kayıt etme yöntemidir. Bu çalışmada BONN veri tabanı kullanılarak EEG verilerinde sağlıklı, nöbet öncesi ve nöbet anı tespit edildi. Çalışmada ilk aşama olarak EEG işaretlerinde ciddi gürültü bulunması nedeniyle ön-ışlem olarak gürültü giderimi gerçekleştirildi. İkinci aşama olarak EEG sinyallerine Ampirik Kip Ayrışımı yöntemi ile İçkin Kip Fonksiyonlarına ayrıştırılarak ve Ayrık Dalgacık Dönüşümü uygulayarak sinyalleri alt-bantlarına ayrıştırarak bu iki frekans uzayı dönüşümü yöntemlerinin ayrı ayrı sınıflandırma aşamasındaki başarımı irdelendi. Sınıflandırma aşamasından önce özellik çıkartımı amacıyla frekans uzayındaki sinyallerin istatistiksel özellikleri çıkartıldı. Bu özellikler kullanılarak Destek Vektör Makineleri (DVM), Karar Ağacı (KA), K-en Yakın Komşuluk (K-YK) sınıflandırıcılarının başarımları analize edildi. BONN veri tabanı üzerinde yapılan deneylerde epilepsi tespiti ve tahmininde en iyi başarımı Ampirik Kip Ayrışımı uygulanılmasının ve sınıflandırıcı olarak da DVM, KA durumunda elde edildiği görüldü. Epilepsi hastalarının mevcut durumdan daha az etkilenmesi amacıyla bu tez kapsamında epilepsinin nöbet tespit ve tahminine yardımcı olacak bir yaklaşım gösterildi.

Anahtar Kelimeler: DVM, EEG, Epilepsi Tahmini, Epilepsi Tespiti, Karar Ağacı.

EPILEPTIC SEIZURE PREDICTION AND DETECTION IN EEG SIGNALS

ABSTRACT

Epilepsy seizure is a neurological disorder that occurs as abnormal electrical discharge in the nerve cells of the brain. Electroencephalography (EEG) is a method of recording this abnormal electrical activity. In this study, healthy, pre-seizure and seizure were determined in EEG data by using BONN database. In the first step of the study, noise removal was performed as a pre-treatment due to the presence of significant noise in the EEG signals. As a second step, the performance of these two frequency space transformation methods in the separate classification stage was examined by separating the signals into sub-bands with Discrete Wavelet Transform, also by separating them into the Intrinsic Mode Functions with Empirical Mode Decomposition (EMD) method to EEG signals. Before the classification, the statistical moments of the signals in the frequency domain were extracted for feature extraction. Using these features, the performances of Support Vector Machines (SVM), Decision Tree, K-nearest Neighborhood (K-NN) classifiers were analyzed. In the experiments performed on the BONN database, the best detection and estimation of epilepsi EEG performance was obtained by applying EMD method and as a classifier SVM and Decision Tree. In this thesis, an approach to assist in the seizure detection and prediction of epilepsy is shown in order to less affect the current situation of epilepsy patients.

Keywords: SVM, EEG, Epilepsy Detection, Epilepsy Prediction, Decision Tree.

GİRİŞ

Elektroansefalografi (EEG), beyin hücreleri tarafından üretilen elektriksel faaliyetlerin dalgalar şeklinde bilgisayar ortamına aktarılması işlemidir. Beyne ait bu elektriksel faaliyetler çeşitli sebeplerle bozulabilmektedir. Epilepsi (sara hastalığı) bu tür durumların başında gelmektedir. Beyin dalgaları EEG'yi oluşturmaktadır, bu dalgalar değerlendirilerek bu hastalık hakkında rahatlıkla bilgi edinilebilir. Epilepsi hastalığı tespit edilirken ya da hangi tip epilepsi hastalığı olduğuna karar verilirken kullanılan yöntem EEG'dir. Fakat bu dalga şekillerini gözle yorum yapmak oldukça güç bir durum olmaktadır. Bu yüzden bu alanda çalışan uzman kişilere ihtiyaç duyulmaktadır. Fakat bu uzmanlık alanına sahip kişileri her yerde bulmak ya da onlara ulaşmak kolay olmadığı için bu kayıtlar hakkında yorum yapabilecek otomatik sistemlere ihtiyaç duyulmaktadır. Bununla ilgili pek çok yöntem geliştirilmiştir.

Bu çalışma da ise özellik çıkartımı için Ayrık Dalgacık Dönüşümü ve Ampirik Kip Ayrışımı yaklaşımları kullanılarak frekans uzayına geçilir ve elde edilen veriden istatistiksel özellikler çıkartılıp sınıflandırma içinse DVM, K-YK ve KA algoritmaları kullanılarak en iyi sistem oluşturulmaktadır. EEG verilerinin sınıflandırılması için istatistiksel işlemler yapılarak öznitelik vektörlerinin boyutları azaltılmıştır. Elde edilen bu vektörlere farklı sınıflandırıcılar denenerek sınıflandırma performansları incelenmiştir.

1. ELEKTROENSEFALOGRAFİ (EEG)

EEG'nin tarihsel gelişimi 1875'lerde hayvanlar üzerinde yapılan deneylerle başlamıştır. Bu deneyler de beyne ait bir takım elektriksel faaliyetlerin varlığı hakkında tartışmalar başlamıştır. İnsan beynine ilk kez elektrotlar yerleştirilerek 1929 yılında insan beynine ait elektriksel faaliyetler gözlemlenebilmiştir. Bunu ortaya koyan ise Hans Berger adında bilim insanıdır. Beynimiz şiddeti düşük olmakla birlikte elektrik akımı üretme ve bu akımlara dayanan dalgaları yayma yeteneğine sahiptir. Bu dalgaların bilgisayarlar ortamına aktarılması işlemi de Elektroensefalografi (EEG) olarak tanımlanmaktadır. Başka bir deyişle insan beynine ait sinyallerin makineler tarafından kayda dökülmesi işlemidir.

Beynin işlev sürecinde beynin uyarılmasıyla oluşan elektriksel potansiyel değişiminin kayıt altına alınması işlemine EEG denilmektedir.

Sadece uyanıklık sırasından bu sinyalleri kaydetmek yeterli bir yaklaşım olmamaktadır. İdeal bir yaklaşım tarzı isteniyorsa bu sinyallerin hem uyanırken hem de uyurken alınması gerekmektedir. EEG beynin fonksiyonları hakkında bilgi vermektedir. EEG tetkiklerini almak için bir yaş sınırlanması bulunmamaktadır.

Vücuda her hangi bir elektrik veya farklı bir uyarıcı verilmediği için EEG yan etkisi söz konusu değildir. Fakat EEG sinyalleri değerlendirilirken hastanın yaşı, uyku ya da uyanık olma durumu vücuda herhangi bir ilaç alımı olup olmadığı gibi durumlar dikkate alınmalıdır. Tutarlı sonuçlar elde edilmesi açısından bu durumlar büyük önem arz etmektedir.

Kayıt için elektrotlar yerleştirilirken uluslararası geçerliliğe sahip 10-20 sistemi dikkate alınır. Bu noktalara ait kayıtlar montaj olarak isimlendirilen bağlantılar vasıtasıyla değerlendirmeye alınır.

EEG karmaşık bir şekle sahip olduğundan dolayı bu konu da yorum yapmak oldukça güçtür. Kafanın üst kısmından alınan EEG'lerin genliği tepeden tepeye 1–100 μV ve frekans bandı 0,5–100 Hz'dir.

EEG işaretleri periyodik değildir, genlik, faz ve frekansları sürekli değişim gösteren bir yapıya sahiptir. O yüzden anlamlı bir veri seti elde edilmek isteniyorsa bu konu da uzun süreli gözlemlere ihtiyaç duyulmaktadır. Kişinin zihnine bağlı olarak değişim göstermektedir EEG frekansları. Bu yüzden kapsadığı frekanslar bakımından EEG değerlendirilmesi yapılır. EEG işaretlerine ait bantlar ve frekans aralıkları Tablo 1.1'de verilmiştir.

Tablo 1.1. EEG işaretlerinin kapsadıkları bantlar

Bant ismi	Sembol	Frekans aralığı
Delta	δ	0,5 – 3,5 Hz
Teta	θ	4 – 7 Hz
Alfa	α	8 – 12 Hz
Beta	β	12 – 22 Hz
Gama	γ	22 – 30 Hz

Delta Dalgaları: 3,5 Hz'in altındaki beyin dalgalarıdır. Süt çocuklarında ve ağır organik beyin hastalıklarında görülür. Genlikleri, 100 μV (p-p)'den küçüktür.

Teta Dalgaları: 4–7 Hz arasındaki dalgalarıdır. Özellikle, çocuklarda parietal ve temporal bölgelerde ortaya çıkarlar. Yetişkinlerde de, duygusal gerginlik, düş kırıklığı durumlarında ortaya çıkarlar. Genlikleri, 100 μV (p-p)'den küçüktür.

Alfa Dalgaları: 8–12 Hz arasındaki beyin dalgalarıdır. Uyanık, normal ve sakin kimselerde görülür. Genlikleri 50 μV kadardır. Uyku durumunda yok olurlar.

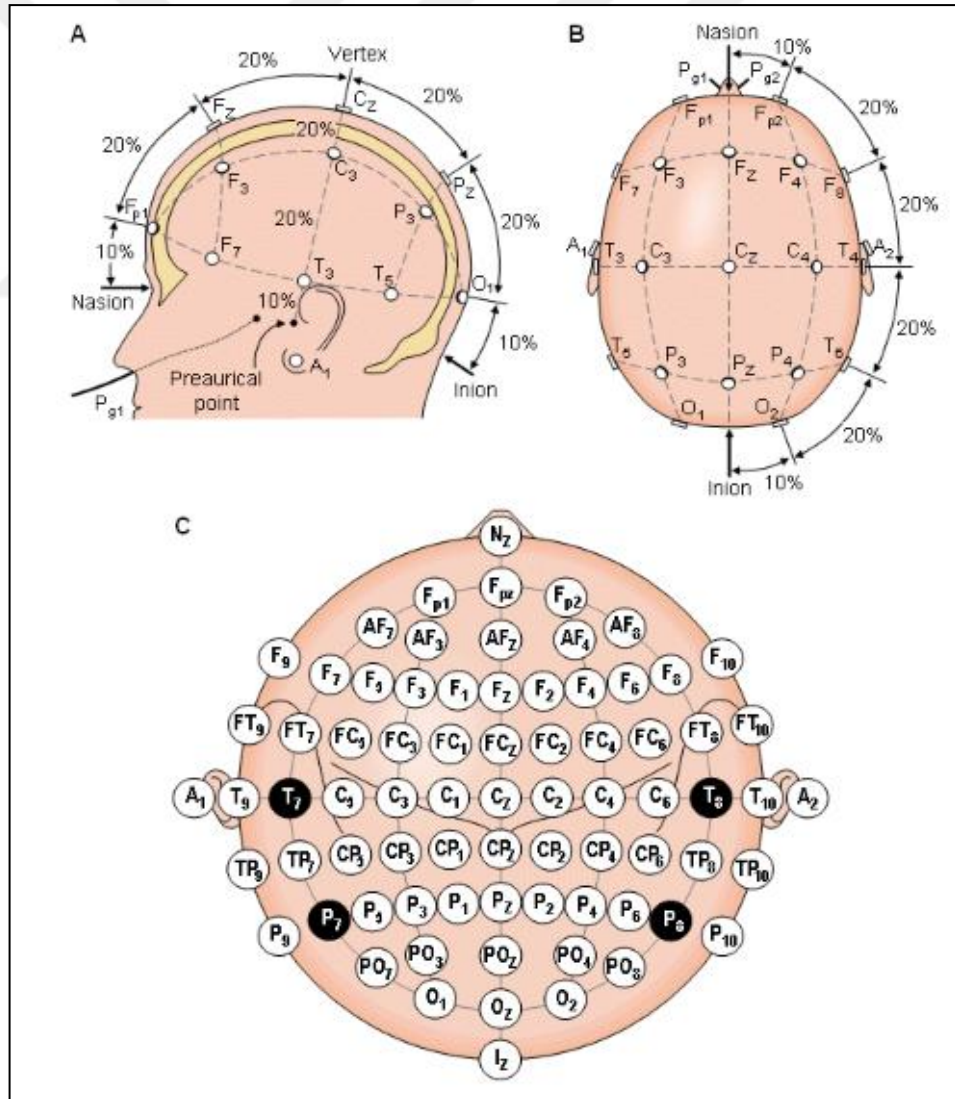
Beta Dalgaları: Frekansları 12 Hz'in üzerindeki beyin dalgalarıdır. 25 Hz'e ve nadir hallerde de 50Hz'e kadar uzanırlar. Beta-I ve Beta-II diye ikiye ayrılırlar. Beta-II dalgalarının frekansı, Beta I'in iki katıdır ve α dalgalarında olduğu gibi zihinsel aktivitenin artması ile ortadan kalkarlar ve yerlerine düşük genlikli asenkron işaretler oluşur. Beta-II dalgaları, merkezi sinir sisteminin kuvvetli aktivasyonunda ve gerginlik hallerinde ortaya çıkar.

Gama Dalgaları: Bazı araştırmacılar tarafından kullanılmaktadır. Bazı araştırmacılarda bu dalgaların yerine, beta-II dalgalarını kullanmaktadır. Genlikleri, 2

4 $\mu\text{V}(\text{p-p})$ 'den daha küçüktür. Kafanın merkezinde, genlikleri daha büyüktür. Uykunun karakteristik belirtisini taşırlar [1].

1.1. Elektroensefalogram İşaretlerinin Ölçüm Yöntemi

EEG işaretlerinin ölçülmesinde bilgisayar ortamından oldukça faydalanılmaktadır. Beyin dalgaları ve bilgisayar etkileşimi ölçüm tekniğini temeli haline gelmiştir. Bu işlem için beyne ait bu elektriksel aktiviteler kafa yüzeyine yerleştirilen metal elektrotlar vasıtasıyla yapılmaktadır. İçerik ve amaç ölçüm elektrotlarının sayısı üzerinde etkili olsa da ölçümler elektrot yerleştirme düzeninde genellikle uluslararası 10-20 sistemi denilen sistem kullanılmaktadır. Şekil 1.1'de uluslararası 10-20 elektrot yerleşim sistemi gösterilmiştir.



Şekil 1.1. Elektrot yerleştirme sistemi [24]

10-20 sisteminde baş için 4 standart nokta ele alınmaktadır. Sırasıyla bunlar burun, başın arka kısmı, sol ve sağ kulak arkaları olmaktadır. Elektrotlar burun ve baş arasına %10-20-20-20-10 olacak şekilde bölünerek yerleştirilir. Sistemin 10-20 sistemi olarak adlandırılmasının sebebi de buradan gelmektedir [1].

1.2. EEG'nin Kullanım Alanları

EEG sinyallerinin başlıca kullanım alanları aşağıda verilmektedir;

1. Nöroloji: EMG, hastanın beyin patolojisi hakkında bilgi vermek amacıyla bu alanda kullanılır.
2. Beyin Cerrahisi: Beyin ameliyatında çıkarılacak tümör vs. dokuların yerlerini belirlemek maksadı ile kullanılır.
3. Anestezi: Anestezi anında hastaya ait anestezi seviyesinin belirlenmesinde kullanılır.
4. Pediatri: Yeni doğan bebeklerin duyma ve görme durumları hakkında yorum yapılmasında kullanılır.
5. Psikiyatri: Zihinsel bir bozukluk teşhisi için beyin hastalığı olup olmadığı hakkında bilgi edinmek maksadı ile kullanılır [1].

1.3. Epilepsi

Epilepsi, beyni etkileyen pek çok faktöre bağlı olmakla birlikte beyin hücrelerinin boşalması sonucunda ortaya çıkan belirtilerin tümü şeklinde tanımlanabilir. EEG, epilepsi tanısı açısından oldukça önemlidir. Tanısı gözlemlenmesi ve tedavisi hakkında bilgi vermektedir. Epilepsi, bir grup beyin hücresinin aşırı boşalmasına bağlı, mental ve motor fonksiyonların geçici olarak kaybolduğu, epileptik krizler ile kendini gösteren nörolojik bir fonksiyon bozukluğudur. Epilepsi krizleri sırasında EEG'de epileptik aktivitenin saptanması, epilepsinin diğer epilepsi benzeri kriz oluşturan hastalıklar ile ayırıcı tanısının yapılabilmesini sağlar. Epileptik krizler sırasında EEG'de yüksek genlikli, tekrarlayıcı, birbiri ardı sıra gelen dalga şekilleri gözlenir. Krizler arası (interictal) dönemlerde ise diken, diken-dalga kompleksi (spike wave), keskin dalga (sharp wave) gibi kısa süreli dalga tipleri kaydedilir [1].

2. LİTERATÜR TARAMASI

EEG verileriyle ilgili çalışmaların tarihçesi 1940'lı yıllara dayanmaktadır. Konuyla ilgili günümüze kadar pek çok çalışma mevcuttur.

EEG sinyallerinin özelliklerini incelemekle başlayan çalışmaları zaman ve frekans uzayına dayalı yöntemler ile EEG sinyallerinin analizi takip etmiştir.

Smith'in yaptığı bir çalışma da EEG'de ani sapsmaları tespit etmek için bir yöntem ortaya koymuştur. Çalışmada dik uçlar (spike) özelliklerini ortaya çıkarmak için özel bir detektör kullanılmıştır [2].

Jones ve diğ. çalışma da çok kanallı bir sistem kullanarak EEG'deki epileptik şekillerin tespiti hakkındadır. Özellik çıkarma aşamasında, her EEG kanalındaki süre, genlik, kesinlik gibi parametrelere dayanarak EEG'deki ani değişimlerini incelemişlerdir [3].

EEG sinyallerini farklı metotlarla incelendikten sonra özellik çıkartma ve sınıflandırma aşamasına geçilmiştir.

Park ve diğ. Dalgacık Dönüşümü (DD) ve Yapay Sinir Ağı (YSA) kullanarak çok kanallı bir epileptik nöbet algılama algoritması üzerine çalışma yapmışlardır. Dik uçların özelliği olarak alfa, beta dalgalarına ait dalgacık katsayıları kullanılmıştır. YSA kullanılarak öğrenme sağlanmış ve oluşturduğu bu YSA modelinin en iyi sonuç verdiği sonucuna ulaşmışlardır. EEG ile epileptik nöbetin tespiti ve tahmini için pek çok farklı çalışma ve yöntem mevcuttur [4].

Srinivasan ve diğ. yaptıkları çalışma elman ağı olarak bilinen özel bir sinir ağı kullanarak, epileptik saptama için otomatik bir teşhis yöntemi üzerinedir. Deneyler EEG sinyalinin frekans-alan özellikleri yanı sıra zaman-alanı özellikleri kullanılarak gerçekleştirilmektedir. Deneysel sonuçlar, elman ağının, iki ve daha fazla giriş özelliğine sahip diğer sinir ağı türlerinin kullanılmasıyla elde edilen sonuçlardan

daha iyi olan tek bir giriş verisi ile epileptik algılama doğruluğu oranını %99 olarak verdiğini göstermektedir [5].

Panda ve diğ. çalışmalarında DD ve DVM kullanarak epileptik ataklar % 91,2 başarımla tespit edilmiştir [6].

Murugavel ve diğ. DD, Lyapunov üsteli ve DVM kullanarak epileptik ataklar %96 oranında tespit etmektedir [7].

Alam ve Bhuyian çalışmalarında varyans, çarpıklık ve basıklık gibi yüksek dereceli istatistiksel anların kullanılmasına dayanan bir nöbet saptama yöntemlerini ele almışlardır. Bu anların EEG sinyallerini ayırt etme kabiliyetini araştırmak için AKA alanında kapsamlı bir analiz yapılmıştır. Bu anlar, bir YSA kullanarak EEG sinyallerini sınıflandırmak için özellikler olarak kullanılmıştır. Önerilen yöntemin, tüm durumlar için başarılı sonuçlar sağlayabildiği gösterilmiştir [8].

Mohite ve diğ. çalışmada EEG verilerinden özellikleri çıkarmak için farklı özellik çıkarım yöntemleri uygulanmıştır. Bu yaklaşımlar Fourier dönüşümü, ADD, AKA ve bispectrum analizidir. Özellikler bu analiz edilen EEG sinyallerinden elde edilip sınıflandırıcı eğitimi için kullanılmıştır. Dört yöntemin de sonuçları AKA'nın daha iyi bir varyansa sahip olduğunu ve bu nedenle sağlıklı ve epilepsi EEG'leri sınıflandırmak için AKA'nın diğer yöntemlerden daha etkin olduğunu göstermişlerdir [9].

Hamad ve diğ. çalışmaların da epilepsi tespiti için ADD yöntemine dayanan bu çalışmada EEG sinyalinden on özellik çıkarılmıştır. Bu özellikler entropi, minimum, maksimum, ortalama, medyan, standart sapma, varyans, çarpıklık, enerji ve Bağlı Dalga Enerjisidir (BDE). Epilepsiyi saptamak için sınıflandırma aşamasında bu özellikleri kullanarak iyi bir doğruluk değeri elde etmişlerdir [10].

Tessy ve diğ. yaptıkları çalışma da kamuya açık bir veri seti üzerinde çalışmışlardır. Satır uzunluğu ve enerji olmak üzere iki zaman alanı özelliği kullanılmıştır. Bunlar sırasıyla Birincisi Kuadratik Diskriminant Analizi (KDA), K-YK ve Lineer Diskriminant Analizi (LDA), EEG sinyallerini sınıflandırmak için kullanılıp sınıflandırma performanslarını değerlendirmiştir. Üç sınıflandırıcının sonuçlarını

karşılaştığında K-YK sınıflandırıcısının diğer iki sınıflandırıcıdan daha iyi sonuçlar verdiği sonucuna ulaşılmıştır [11].

Das ve Mudoji çalışmalarında AKA alanındaki EEG sinyallerinden epileptik nöbetin tespiti yöntemini ele almışlardır. Varyans, çarpıklık ve basıklık gibi istatistiksel anlar ve örnek entropi ve yaklaşık entropi gibi doğrusal olmayan ölçülerin İçkin Kip Fonksiyonları (İKF) AKA yöntemi sayesinde elde edilmiştir. Sınıflandırma için hesaplanan özellikler DVM yöntemi ile eğitilip, incelenmiştir. Veri setinde sağlıklı, interiktal ve iktal olmak üzere üç durum göz önünde bulundurulmuştur. Önerilen yöntem, sağlıklı ve iktal, interiktal ve iktal EEG sinyallerinin sınıflandırılmasında iyi sonuçlar vermiştir [12].

Yol ve diğ. yaptıkları çalışma da epileptik nöbetlerin saptanması için AKA kullanarak EEG sinyallerinin sınıflandırılması amaçlanmıştır. Bunun için çalışmada Bonn Üniversitesi Epileptoloji Anabilim Dalı'nın veri tabanından alınmış olan veri seti kullanılmıştır. Sağlıklı bireylerden ve hastalardan alınan bu kayıtlar AKA yöntemi ile İKF'lere ayrıştırılmaktadır. Tsallis Entropy, Renyi Entropy, Relative Entropy ve Coherence yöntemlerine dayanılarak özellik vektörleri elde edilmiştir. Bu özellikler daha sonra K-YK, Doğrusal Ayrıştırma Analizi (LDA) ve Naive Bayes Sınıflandırıcısı (NBS) ile sınıflandırılmıştır. Çalışmanın sonunda sağlıklı ve hasta EEG verileri arasında önemli farklılıklar tespit edilmiştir [13].

Daoud ve diğ. yaptıkları çalışma da EEG kayıtlarından otomatik epileptik nöbet tespiti için sunulan yöntemleri ele almışlardır. Önerilen yöntemde, özellik çıkarımı için AKA ve sınıflandırma için Derin Sinir Ağları (DSA) kullanılmıştır. Çok katmanlı algılayıcı, özelliklerin boyutunu azaltarak sağlıklı ve nöbet durumlarını sınıflandırmak için kullanılmıştır. Bu yöntem kullanılarak elde edilen sınıflandırma doğruluğu %100'dir. Derin Evrişimli Sinir Ağları (DESA), ikinci sınıflandırma yönteminde, çok sınıflı sınıflandırma görevinde yüksek bir doğruluk elde etmek için kullanılır. Sağlıklı, interiktal ve iktal olgular arasında sınıflandırma için DESA kullanılarak elde edilen sınıflandırma doğruluğunu %98,6 olarak bulmuşlardır. Sonuçlar, önerilen yaklaşımın (sağlıklı-iktal) ve (sağlıklı-iktal-interiktal) EEG sinyallerinin yüksek doğrulukla sınıflandırmada etkili olduğunu göstermektedir [14].

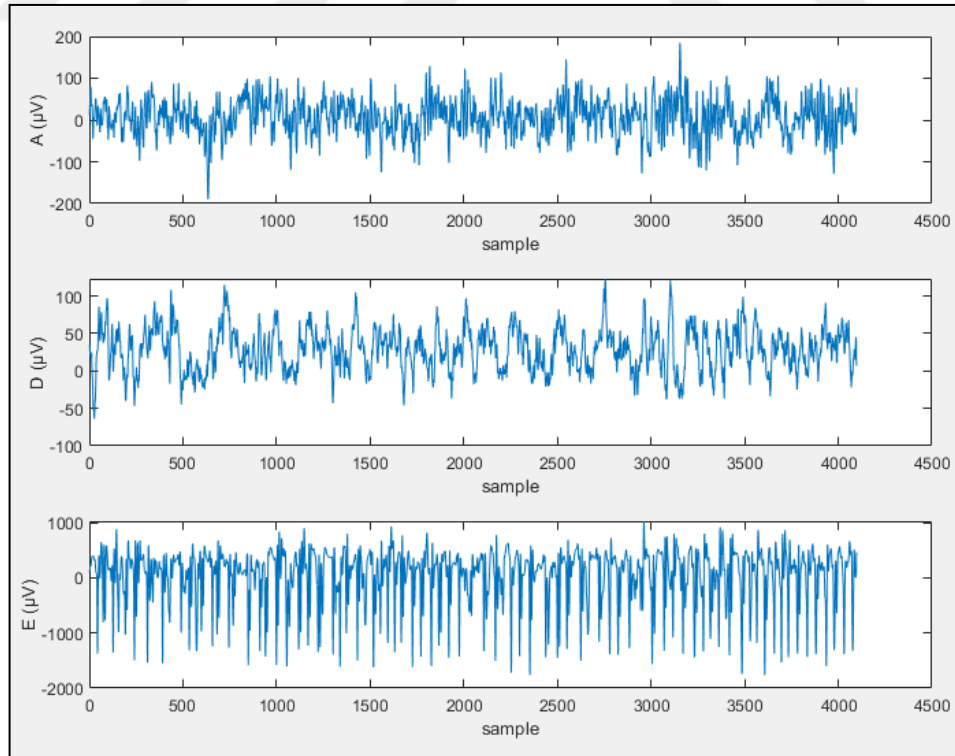
Bu tez çalışmasında ise öncelikle EEG işaretlerinde ciddi gürültü bulunması nedeniyle ön-işlem olarak gürültü giderme işlemi yapılmıştır. EEG sinyallerine Ampirik Kip Ayrışımı yöntemi ile İçkin Kip Fonksiyonlarına ayrıştırılarak ve Ayrık Dalgacık Dönüşümü uygulayarak sinyalleri alt-bantlarına ayrıştırarak bu iki frekans uzayı dönüşümü yöntemlerinin ayrı ayrı sınıflandırma aşamasındaki başarımı değerlendirilmiştir. EEG işaretleri üzerinden sağlıklı, nöbet öncesi ve nöbet anı tespit edilmiştir. Sınıflandırıcı olarak DVM, KA, K-YK algoritmaları kullanılarak en iyi performansı sergileyen yöntem belirlenmeye çalışılmıştır.



3. MATERYAL VE YÖNTEM

3.1. Materyal

Çalışmada Bonn Üniversitesine ait olan veri tabanı kullanılmaktadır. Veri seti A, B, C, D, E olarak adlandırılan 5 kümeden oluşmaktadır. Her veri kümesi 100 tek kanallı EEG verisi içermektedir. Sinyallerin örnekleme frekansı 173,61 Hz olduğundan örnekleme frekansının yarısı 86,81 Hz'dir. A ve B veri kümeleri gözleri açık ve kapalıyken beş sağlıklı gönüllüden toplanan EEG bölümleridir (tüm beyin bölgeleri). Küme C ve D, nöbet öncesi dönemlerini içeren EEG verileridir (epileptik bölge ve epileptik bölgenin karşısı). E kümesi de nöbet sırasındaki döneme ait EEG kayıtlarıdır (epileptik bölge) [15]. Sağlıklı, nöbet öncesi ve nöbet anına ait EEG sinyaller ise Şekil 3.1'de gösterilmiştir. Bu şekilde yatay eksen örnek sayısını (sample), dikey eksen ise örneklerin genliklerini (μV) göstermektedir.

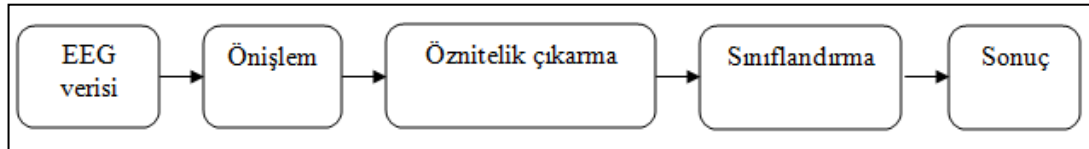


Şekil 3.1. EEG sinyalleri: A- sağlıklı; D- nöbet öncesi; E- nöbet anı verileri

EEG sinyali 0-60 Hz frekans aralığında yayılmaktadır. Dolayısıyla 60 Hz'den büyük frekanslar ciddi gürültü olarak karakterize edilir. Bu nedenle, çalışmada önışlem için Notch filtresi uygulanmaktadır. Bu filtre bant durduran bir filtredir ve 50 Hz'lik frekansı bastırmaktadır. Ayrıca 50 Hz'lik Notch filtresi şebekeden gelen gürültüyü bastırmaktadır.

3.2. Özellik Çıkartımı

Bu çalışmada, otomatik değerlendirmeden faydalanıp EEG kayıtlarına dayanarak epilepsi teşhisinin yapılması amaçlanmıştır. Tıbben karar vermek ve doğru teşhis yapmak önemlidir. Bu yüzden eldeki veri seti, veri özelliklerinin çıkarılması, çıkarılan özelliklerden faydalanıp veriyi analiz etmek büyük önem arz etmektedir. Uzmanlar istatistiki tekniklerden faydalanıp veriyi işleme hazır hale getirmektedirler. Veri seti büyüdükçe bu nu işleme sokmak da o derece güçleşmektedir o yüzden makine öğrenmesine ihtiyaç duyulmaktadır. Bu süreçte bilgisayar destekli veri sınıflamasına ihtiyaç duyulmaktadır. Günümüzde bilgisayar destekli karar verme uygulamaları hızla artmaktadır. Hastalık teşhisleri, veri sınıflama işlemleri arasında yerini almaktadır. Teşhis sistemlerinin kullandığı mekanizma aşamaları önışleme, öznitelik çıkarma ve sınıflama işlemleridir. Öznitelik çıkarma, şekle dayanan öznitelik vektörünün elde edilmesidir, öznitelik seçme ise sınıflama işlemi açısından en belirleyici özelliklere karar verip öznitelik vektörünün boyutunu azaltmak amaçlanmıştır. Sınıflandırma ise teşhis sisteminin son aşamasıdır. Kullanılan algoritmaya bağlı olarak giriş öznitelik vektörleri incelenip sınıflandırma sonucu değerlendirilir. Çalışmaya ait algoritma Şekil 3.2'de verilmiştir.



Şekil 3.2.Epilepsi nöbet tahmini ve tespiti için blok şeması

Şekil 3.2'de EEG işaretlerinin analizinde ADD, AKA ve istatistiksel momentleri baz alınarak yapılmıştır. Elde edilen öznitelik vektörleri DVM, K-YK, KA sınıflandırıcıları uygulanarak sağlıklı, nöbet öncesi ve nöbet anı durumları tespit edilmiştir.

3.2.1. Ayrık dalgacık dönüşümü (ADD)

Sürekli olarak skalalandırılmayan ve ötelenemeyen sadece ayrık adımlarla skalalandırılan ve ötelenebilen fonksiyonlar Ayrık baz fonksiyonları olarak isimlendirilir. Parçalı sürekli olan bu fonksiyonlar Ayrık Dalgacık olarak isimlendirilirler. Ayrık baz fonksiyonları Denklem (3.1)'deki şekilde ifade edilmiştir;

$$\Psi_{j,k} = 1/a_0^j \Psi \left\{ (t - kb_0 a_0^j) / a_0^j \right\} \quad (3.1)$$

Fonksiyonlar da tanımlanan j ve k , tamsayı olup, a_0 ise $a_0 > 1$ aralığından olmak şartıyla sabit genişleme adımını temsil etmektedir. Ayrıklaştırma ise zaman-skala uzayının ayrık aralıkla örneklenmesi amacıyla yapılmaktadır. Genelde $a_0 = 2$ olarak alınmaktadır. Böylelikle frekans ekseninin örneklenmesi, örnekleme aralığının ikinin katları olarak arttığı dyadic (ikinci) örnekleme karşılık düşer. Aynı şekilde $b_0 = 1$ alınarak; zaman ekseninin dyadic (ikinci) örnekleme, yani öteleme parametresinin ($u = kb_0 a_0^j$) ikinin katları olarak değişmesi elde edilir. Bu seçimler bilgisayarlı hesaplamalar açısından oldukça verimli bir performans elde edilmesine sebebiyet vermektedir. Süreklilik arz eden sinyale ait, ayrık baz fonksiyonları kullanılarak DD'nün alınması ile bir dizi Dalgacık katsayısı elde edilmektedir. Bu işlem literatürde "Dalgacık Dizi (Seri) Ayrıştırması" olarak adlandırılmaktadır. Fakat bu işleme sinyali geri çatması her zaman mümkün değildir. Sinyalin geri çatılabilmesi için gerekli olan şart aşağıda ki denklemlerle verilmiştir. Elde edilen Dalgacık katsayılarının enerjisinin iki pozitif sınır arasında olmasıdır. $\|f\|^2$ sinyalin enerjisi, $A > 0$, $B < \infty$ ve A, B sinyalden bağımsız olmak üzere aşağıdaki Denklem (3.2)'de ifade edilmiştir;

$$A \|f\|^2 \leq \sum_{j,k} \left| \langle f, \Psi_{j,k} \rangle \right|^2 \leq B \|f\|^2 \quad (3.2)$$

Yukarıda verilen bu şart sağlandığında; baz fonksiyonlarının hepsi birden ($\Psi_{j,k}$, $j, k \in \mathbb{Z}$) sınırları A ve B olan "çerçeve" (frame) olarak adlandırılır. Sürekli Dalgacık Dönüşümlerinde oldukça fazla bilgidan arındırılması maksadı ile ayrık baz fonksiyonlarının kendi aralarında orthonormal olarak sınıflandırılması gerekir. Bu da sadece ayrık dalgacıklar ile sağlanabilmektedir. Ayrıca başka bir açıdan değerlendirilirse ana dalgacık fonksiyonun özel olarak seçilmesiyle, ayrık baz fonksiyonlarının kendi aralarında orthogonal olması da sağlanabilir. İki vektörün

veya iki fonksiyonun iç çarpımları sıfıra eşitse; iki vektör veya iki fonksiyonun orthogonal olduğu söylenebilir. Ayrık dalgacıkların orthonormal ve orthogonal olması için gerekli olan denklemler aşağıda Denklem (3.3)'te verilmiştir;

$$\int \Psi_{j,k}(t) \Psi_{m,n}^*(t) dt = \begin{cases} 1, & j=m, k=n \\ 0, & j \neq m, k \neq n \end{cases} \quad (3.3)$$

Seçilen herhangi bir sinyal, aşağıda verilen Denklem (3.4)'te gösterildiği gibi DD katsayıları ile ağırlıklı orthogonal dalgacık baz fonksiyonlarının toplamı olarak tekrar elde edilebilir yani geri çatılabilir. Bu ifade ayrık baz fonksiyonları kullanılması durumunda Ters Dalgacık Dönüşümü olarak isimlendirilir;

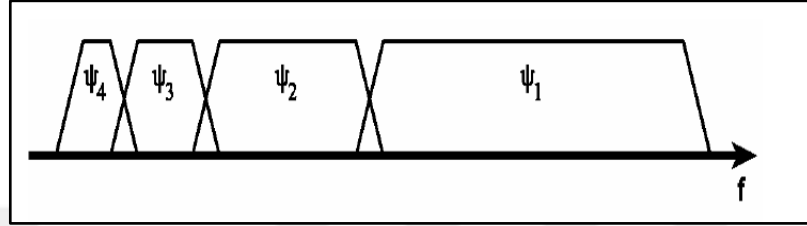
$$f(t) = \sum_{j,k} \gamma(j,k) \Psi_{j,k}(t) \quad (3.4)$$

Baz fonksiyonlarının orthogonal olması sinyale ait zaman frekans gösterimi açısından şart değildir. Fazla bilgiden arındırılması gerektiğinde bu şart konulmaktadır.

Gürültüye karşı duyarlılığı azaltmak için bu bilgilere bazen gereksinim olabilir. Böyle durumlarda gürültüye karşı olan duyarlılığı azaltmak adına fayda sağlamaktadır ve böylece dönüşüm sinyalinin ötelenmesi ile değişmemesi durumu açısından fayda sağlar. Fakat ayrık baz fonksiyonları ile DD'de, bir sinyalin zamanda ötelenmiş halinin dönüşümü, kendi dönüşümünün sinyalin zamanda ötelenme miktarı kadar ötelenmişine eşit olmamaktadır. Bu ayrık dalgacıkların kullanılmasının yol açtığı tek dezavantaj olarak sayılabilir. Unutulmaması gereken önemli bir nokta ayrık baz fonksiyonları kullanılarak hala sürekli sinyallerin DD'nün alınıyor olmasıdır.

Baz fonksiyonlarının, analizi yapılan sinyalin zaman dönemindeki süresi yani diğer bir deyişle boyu kadar ötelenmesi gereklidir. Dalgacık için gerekli olan üst sınır böyle belirlenir fakat sinyalin analizi için kaç skala değerine ihtiyacımız olduğuna başka bir deyişle Dalgacık için bir alt sınıra ihtiyaç duyulmaktadır. Bu sorunun çözümü için DD'ne farklı açıdan değerlendirilebilir. Baz fonksiyonunun zaman döneminde 2 çarpanı ile sıkıştırılması, baz fonksiyonun spektrumunun 2 çarpanı ile genişletilmesine ve tüm frekans bileşenlerinin 2 çarpanı ile frekans döneminde kaydırılmasına karşılık gelmektedir. Bu özellik dikkate alınarak, zaman döneminde ötelenmiş baz fonksiyonları ile sinyal nasıl tümüyle kaplanıyorsa aynı yolla frekans

döneminde sinyalin sonlu uzunluktaki spektrumu genişletilmiş baz fonksiyonlarının spektrumları ile kaplanabilmektedir. Frekans dönemlerinde sinyalin genişletilmiş baz fonksiyonlarının spektrumları ile en iyi şekilde kaplanması, verildiği gibi, genişletilmiş baz fonksiyonlarının spektrumlarının birbirleriyle kesişecek şekilde baz fonksiyonlarının ayarlanması ile elde edilir [16]. Şekil 3.3'te zamana bağlı ana DD fonksiyonu yer almaktadır.



Şekil 3.3. Zamana bağlı ana DD'nün skalandırılması ile elde edilen temel fonksiyonlarının spektrumları (2 çarpanı ile) [17]

Tek bir baz fonksiyonu, bir bant geçiren filtre olarak düşünülürse, skalalandırılmış baz fonksiyonlarının dizisi ise, bant geçiren filtreler kümesi olarak görülebilir. Eğer tüm baz fonksiyonlarının spektrumlarının merkez frekansı ve spektrumun genişliği arasındaki orana bakılırsa, hepsi için aynı olduğu gözlemlenir. Bu orana “sadakat faktörü (Q)” denir. Tüm baz fonksiyonları ise “Sabit Q faktörlü filtre kümesi” olarak adlandırılır [16].

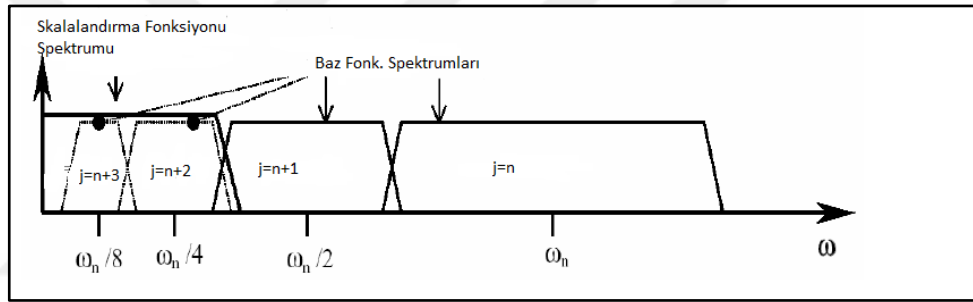
Eğer sinyalin enerjisi sonsuz ise, sinyalin spektrumunun veya zaman uzyında sinyalin, baz fonksiyonları ile kaplanması mümkün olmamaktadır. Öyleyse aşağıdaki Denklem (3.5) göz önüne alınıp ifade edildiği gibi sinyalin enerjisinin sonlu olması gerekmektedir;

$$\int |f(t)|^2 dt < \infty \quad (3.5)$$

Fakat ana dalgacık fonksiyonun zaman döneminde 2 çarpanı ile genişletilmesi sonucu elde edilen fonksiyonun spektrumunun bant genişliği, ana dalgacık fonksiyonunun spektrumunun bant genişliğinin yarısını ifade etmektedir. Yani zaman döneminde genişletme, frekans döneminde aynı oranda sıkıştırmaya karşılık gelmektedir. Bu şekilde sıfıra doğru yaklaşarak sinyal spektrumunun sonlu sayıda baz fonksiyonu ile kaplanması oldukça güçtür. Çünkü baz fonksiyonları ana dalgacık fonksiyonundan 2 çarpanı ile genişletilerek elde edileceğinden, her seferinde elde

edilen baz fonksiyonun spektrumunun bant genişliği bir öncekinin yarısına eşit olmaktadır. Bu durumda sonlu sinyal spektrumunun kaplanması için sonsuz sayıda baz fonksiyonuna ihtiyaç duymaktadır [16].

Sözü edilen sorun verildiği gibi sıfıra doğru sonsuz sayıdaki baz fonksiyonunun spektrumu yerine bu alanı kaplayacak bir alçak geçiren filtre spektrumunun yerleştirilmesiyle çözümlenebilir. Bu çözüm ilk defa 1989'da Mallat tarafından skalalandırma fonksiyonunun tanımı ile gerçekleştirilmiştir. Bu fonksiyonun spektrumu gerekli olan alçak geçiren filtre spektrumuna karşılık gelir. Skalalandırma fonksiyonunun spektrumu, alçak geçiren özelliğinden dolayı ortalama filtresi olarak da adlandırılmaktadır [16]. Sinyal spektrumunun sonlu sayıda spektrum ile kaplanmasının sağlanmasına dair görsel Şekil 3.4 ile gösterilmiştir.



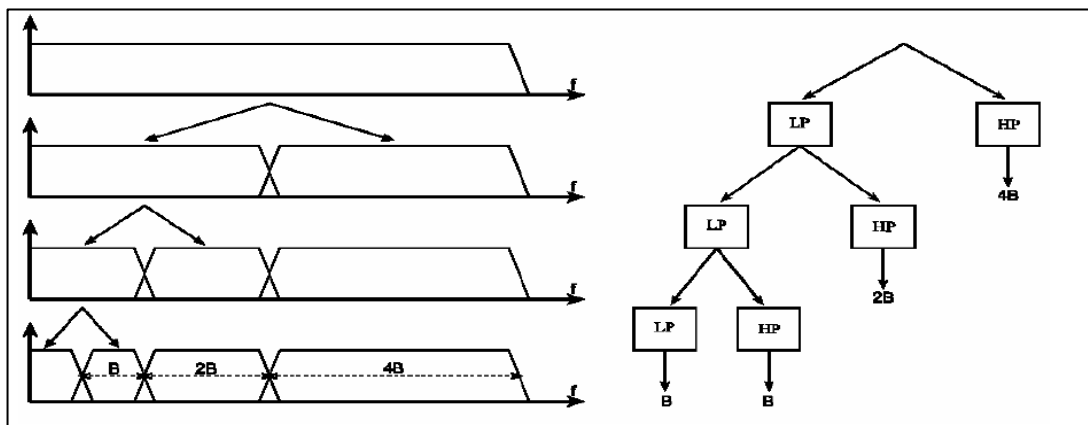
Şekil 3.4. Sinyal spektrumunun sonlu sayıda spektrum ile kaplanmasının sağlanması [17]

Mallat tarafından ortaya konulan skalalandırma fonksiyonu ile sonsuz sayıda baz fonksiyonu gereksinimi ortadan kaldırılmıştır. Böylece Dalgacık için bir alt sınır belirlenmiştir. Fakat bir kısım baz fonksiyonunun yerine skalalandırma fonksiyonunun kullanılması bilgi kaybına neden olmaktadır. Bilgi kaybı, sinyalin geri çatılmasını yani DD'nden tekrar elde edilebilmesini engellemeyecek düzeyde olduğu sürece sinyal gösterimi açısından bilgi kaybı olmadığı düşünülmektedir. Ancak Dalgacık analizi açısından değerli skala bilgilerin atılması anlamını taşımaktadır. Bu yüzden skalalandırma fonksiyonunun genişliği DD açısından oldukça önemlidir. Skalalandırma fonksiyonu spektrumunun daralması daha fazla Dalgacık katsayısı ve daha fazla skala bilgisi elde edilmesini sağlamaktadır. Fakat Dalgacık katsayılarının miktarı için pratik olarak kontrol edilebilmelerini sağlamak amacıyla bir sınır mevcuttur. Bir baz fonksiyonu bant geçiren filtre ve skalalandırma

fonksiyonu da alçak geçiren filtre olarak kabul edilirse, tüm baz fonksiyonları ve skalalandırma fonksiyonu bir filtre kümesi olarak kabul edilmektedir.

DD bir filtre kümesi olarak göz önüne alınırsa, DD işlemini sinyalin bu filtre kümesinden geçirilmesi olarak düşünülebilir. Her farklı filtrenin çıkışı Dalgacık ve skalalandırma fonksiyonu dönüşüm katsayılarıdır. Bu tarz bir analiz yeni bir fikir değildir ve yıllardır “alt bant kodlaması” adı altında kullanılmaktadır [16].

Alt bant kodlamada gerekli olan filtre kümesi birkaç çeşit yolla kurulabilir. Örneğin, sinyalin spektrumu birçok bant geçiren filtre ile frekans bantlarına bölünebilir. Bu yolla her frekans bandının genişliğinin serbestçe belirlenebilmesi bir avantajdır, fakat her filtrenin ayrı olarak tasarlanması gerektiğinden bu işlem zaman açısından çok kullanışsızdır. Diğer bir yol ise, sinyalin spektrumu alçak geçiren ve yüksek geçiren olmak üzere iki parçaya bölünür. Yüksek geçiren kısmı genelde bizim ilgilendiğimiz detayları içeren kısımdır. Bu durumda iki frekans bandı elde edilmiştir. Eğer istenilen bilgi elde edilebilirse bu işlem burada sonlandırılır. Fakat hala alçak geçiren kısımda ilgilendiğimiz sinyalin detayları yer almaktadır ve bunların incelenmesi amacıyla alçak geçiren kısmı tekrar alçak ve yüksek geçiren olmak üzere iki parçaya bölünebilir. Bu işlem, istenilen bilgilerin elde edilmesine kadar devam edilebilir. Bu yolla iteratif filtre kümesi elde edilmiş olur. Frekans bantlarının sayısı genelde bilgi veya hesaplanabilir güç miktarına göre sınırlıdır. Şekil 3.5’te verilen yöntemde iki filtre tasarlanması yeterli olacaktır, fakat sinyal spektrumunun filtrelerin spektrumları ile kaplanma şekli hep sabittir [16].



Şekil 3.5. İteratif filtre kümesi ile sinyal spektrumunun frekans bantlarına ayrılması [16]

Bu yöntemde sinyal spektrumu ilk kez bir alçak geçiren ve yüksek geçiren filtre ile ikiye bölündüğünde elde edilen frekans bantlarından biri olan yüksek geçiren bant aslında sinyalin bant genişliği tarafından sınırlandırıldığı için bant geçirendir.

Diğer bir deyişle, bu alt bant kodlama analizi sinyalin frekans ekseninde sağa doğru gidildikçe her seferinde solundaki filtre spektrumunun bant genişliğinin iki katı bant genişliğine sahip bir bant geçiren filtre kümesi ve bir alçak geçiren filtreden geçirilmesi ile gerçekleştirilebilir. Şekil 3.5'te sinyalin spektrumu belli frekans bantlarına ayrılmış olur. Gerçekleştirilen işlem ile DD özdeştir. Alçak geçiren filtre spektrumu ise skalalandırma fonksiyonunun spektrumu ile özdeştir [16].

Belirli bir skala değerindeki skalalandırma fonksiyonu, bir sonraki daha küçük skala değerindeki ötelenmiş skalalandırma fonksiyonu cinsinden ifade edilebilir. Daha küçük skala değeri daha fazla detay anlamına gelmektedir. İlk skalalandırma fonksiyonunun DD ile ayrıştırılmasında kullanılan ayrık baz fonksiyonları, bir sonraki skala değerindeki ötelenmiş skalalandırma fonksiyonu cinsinden ifade edilebilir. j seviyesi için skalalandırma fonksiyonu ve Dalgacık fonksiyonu arasındaki iki skala ilişkisinin matematiksel ifadesi Denklem (3.6)'da verilmiştir;

$$\Psi(2^j t) = \sum_k g_{j+1}(k) \phi(2^{j+1} t - k) \quad (3.6)$$

$f(t)$ sinyali, $-\infty$ 'dan $j-1$ skala değerine kadar ana dalgacık fonksiyonunun ötelenmesi ve skalalandırılması ile elde edilen baz fonksiyonları cinsinden ifade edilebilir. Sinyal, Denklem (3.7)'de verildiği gibi j skala değerindeki skalalandırma fonksiyonunun ötelenmiş ve skalalandırılmış versiyonları cinsinde de ifade edilebilir;

$$f(t) = \sum_k \lambda_j(k) \phi(2^j t - k) \quad (3.7)$$

Bu ifadede skala değeri $j-1$ yapılırsa, ayrıntı seviyesinin aynı kalabilmesi için mutlaka bu ifadeye baz fonksiyonlarının eklenmesi gerekir. Bu durumda sinyal Denklem (3.8)'de verildiği şekilde olur;

$$f(t) = \sum_k \lambda_{j-1}(k) \phi(2^{j-1} t - k) + \sum_k \gamma_{j-1}(k) \Psi(2^{j-1} t - k) \quad (3.8)$$

Eğer ayrık skalalandırma fonksiyonları $\phi_{j,k}(t)$ ve ayrık baz fonksiyonları $\Psi_{j,k}(t)$ orthonormal iseler, $\lambda_{j-1}(k)$ ve $\gamma_{j-1}(k)$ katsayıları Denklem (3.9) ve (3.10) ile verilen iç çarpım ifadeleri ile bulunabilir;

$$\lambda_{j-1}(k) = \langle f(t), \phi_{j,k}(t) \rangle \quad (3.9)$$

$$\gamma_{j-1}(k) = \langle f(t), \Psi_{j,k}(t) \rangle \quad (3.10)$$

Bu iç çarpım ifadelerinin integral formunda $\phi_{j,k}(t)$ ve fonksiyonlarının $\Psi_{j,k}(t)$ belirtildiği gibi skalalandırılmış ve ötelenmiş versiyonları kullanılırsa aşağıdaki Denklem (3.11) ve (3.12) ile bağıntılar elde edilir;

$$\lambda_{j-1}(k) = \sum_m h(m-2k) \lambda_j(m) \quad (3.11)$$

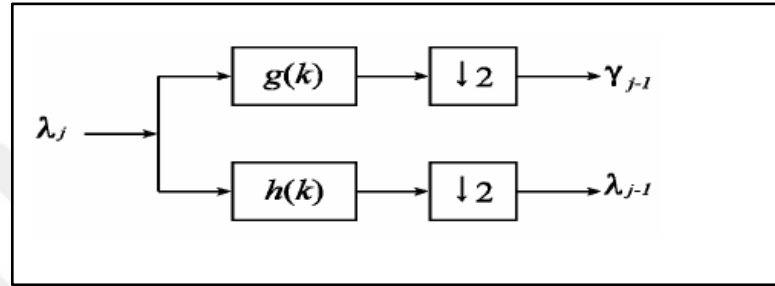
$$\gamma_{j-1}(k) = \sum_m g(m-2k) \lambda_j(m) \quad (3.12)$$

Bu iki ifadeden anlaşıldığı gibi, belli bir skala değerindeki dalgacık ve skalalandırma fonksiyonu katsayıları, bir önceki skala değerindeki skalalandırma fonksiyonu katsayılarının ağırlıklı ortalamasının hesaplanması ile bulunabilir. Elimizdeki sürekli sinyalden belli bir örnekleme frekansında örneklenecek elde edilmiş ayrık sinyal $f(k)$, en büyük skala değerindeki $\lambda(k)$ katsayılarına eşittir.

$\lambda(k)$ katsayılarının sinyalin spektrumunun bölünmesi ile elde edilen alçak geçiren kısımdan geldiği bilindiğine göre; ağırlıklandırma faktörleri $h(k)$, mutlaka bir alçak geçiren filtreyi teşkil etmelidirler. Yani $h(k)$, bir alçak geçiren filtrenin impuls cevabı olmalıdır. $\gamma(k)$ katsayılarının sinyalin spektrumunun bölünmesi ile elde edilen yüksek geçiren kısımdan geldiği bilindiğine göre; $g(k)$, bir yüksek geçiren filtrenin impuls cevabı olmalıdır. Bu durumda yukarıdaki denklemlerde verilen ifadelerin iteratif dijital filtre kümesinin bir basamağını ifade ettikleri açıktır. Bundan dolayı $h(k)$ katsayıları, “Skalalandırma filtresi” ve $g(k)$ katsayıları ise “Dalgacık filtresi” olarak adlandırılır [16].

2 çarpanı nedeniyle, j skala değerindeki $\lambda_j(k)$ katsayılar dizisinin sadece her iki elemanından biri konvolüsyon işlemine katılır. Bu işlem “seyrek örnekleme” olarak adlandırılır. İteratif filtreleme basamaklarının herhangi bir j seviyesinde elde edilen

$\lambda_{j-1}(k)$ ve $\gamma_{j-1}(k)$ katsayı dizilerinin eleman sayıları toplamı, giriş bilgisi katsayı dizisinin eleman sayısına eşittir. Her bir sonraki basamakta örnek sayısı bir öncekinin yarısına düştüğü için iterasyonun sonlandırılması gereken noktası skalalandırma fonksiyonu spektrumunun genişliğini belirler. İterasyon işleminin sonlandırılması için örnek sayısının, skalalandırma filtresi impuls cevabı $h(k)$ veya dalgacık filtresi impuls cevabı $g(k)$ dizilerinden uzunluğu büyük olandan küçük olması gerekir [17]. Şekil 3.6'da iteratif filtreleme gerçekleştirilmesi işlemi gösterilmektedir.



Şekil 3.6. İteratif filtreleme işleminin gerçekleştirilmesi [16]

ADD ile bir sinyal spektrumu, alçak frekanstan yüksek frekanslara doğru peş peşe frekans bantlarına ayrıştırılır. Ayrıştırmanın ilk basamağında işleme tabi tutulan alçak frekans bandı sinyalin spektrumudur. Sinyalin sonlu enerjiye sahip olması gerektiği unutulmamalıdır [16].

3.2.2. Ampirik kip ayrışımı (AKA)

Frekans dönüşüm tekniği olan AKA işaret işlemede oldukça yaygın kullanılmaktadır. Uyarlamalı ve doğrusal olmayan işaret ayrışım yöntemidir. Doğrusal olmayan verilerin ya da işaretlerin işlenmesinde son zamanlarda etkin şekilde kullanılmaktadır. Zaman serisi şeklinde olan işaretleri İçkin Kip Fonksiyonu (İKF) şeklinde işleme alır ya da temsil eder. İçkin kipten kasıt orijinal olan işaretin farklı frekans unsurlarıdır. Bu yöntem kullanılarak işaret içsel mod ve kalanı olarak ayrıştırılır. İçsel mod hesaplama işleminde sınır değerler işaretin sıfır geçiş sayısı ya aynı olmalı ya da aralarındaki farkın bire eşit olması gerekir. Herhangi bir t anında yerel maksimum ve yerel minimum noktalarının birleşimi ile alt ve üst zarfın ortalama değeri sıfıra eşit olmalıdır. İKF ve kalan toplamı ilk işareti verebilmelidir. Bu bilgi kaybının ve hatanın olmadığı anlamını taşımaktadır.

Fourier dönüşümü vb. dönüşümlerde verinin doğrusal ve durağan olduğu düşünülürken AKA'da veri seti hakkında bu varsayım söz konusu değildir. İKF'lerine göre sınıflamaktadır. Özellikle doğrusal ve durağan olmayan problemlerin zaman serilerinde bu yöntem oldukça güçlüdür. Amaç durağan olmayan problemlere ait zaman frekans analizi için anlık frekans değerlerini baz alan ayrışımı yapmaktır. Özyineli sinyal ayrışımı diye tabir edilmesinde sakınca bulunmamaktadır. AKA orijinalde tek boyutlu sinyaller için yapılmış olsa da artık iki boyutlu sinyallere de uyarlanır olmuştur.

AKA sonucunda veriden sonlu sayıda İKF'leri ve bir kalan işaret kısımları elde edilir. İKF'leri yerel anlık frekansları içerir, düşük seviye de olan İKF'leri ise alçak yerel frekansları içermektedir. Tüm İKF'lerinin aşağıda verilen şu iki koşulu sağlaması gerekmektedir;

1- Her İKF sıfır geçişi ve yerel uç noktası sayıları eşit olmalı ve farkları en fazla bire eşit olmalıdır.

2- En büyük ve en küçük noktalardan elde edilen üst ve alt zarf ortalamaları her nokta da sifıra eşit olmalıdır. Bu fonksiyonların eldesi için kullanılan özyinelemeli yaklaşım ise aşağıda verilen gibidir;

- İşaretlere ait yerel uç noktalar belirlenir.
- En büyük ve en küçük noktalara ait ara değerlendirme ile üst zarf $e_{maks}(m,n)$ ve alt zarf $e_{min}(m,n)$ elde edilmektedir.

3- Üst ve alt zarfın ortalaması ile ortalama zarfı Denklem (3.13) ile elde edilir;

$$e_{ort} = e_{maks}(m,n) + e_{min}(m,n) \quad (3.13)$$

4- Giriş işaretlerinden ortalama zarf işaretleri Denklem (3.14) ile çıkarılır;

$$h_{m,n} = giriş(m,n) - e_{ort}(m,n) \quad (3.14)$$

5- Ortalama zarf işaretinin İKF'inin durma kriterinde sifıra yeterince yakın olmayı sağlayıp sağlamadığı kontrol edilir Denklem (3.15);

$$\sum_m^W \sum_n^H |e_{ort}(m,n)| / (W \cdot H) < \tau \quad (3.15)$$

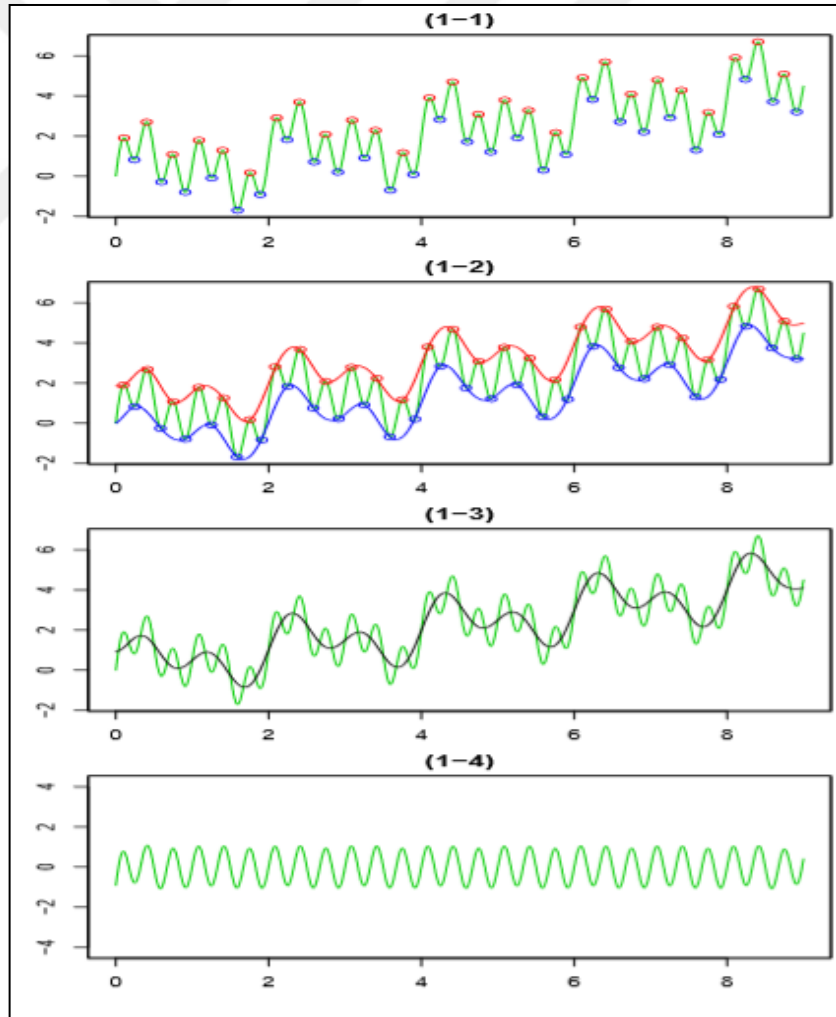
Burada W ve H uzamsal boyutları ve τ ise ampirik olarak belirlenen sifıra yakın eşik değeri (0,2-0,3) göstermektedir.

6- Durma kriteri sağlanmazsa 1. adımdan $h_{m,n}$ yeni giriş sinyali olarak devam edilir. Durma kriteri sağlanırsa $h_{m,n}$ yeni İKF olarak alınır.

7- Bir sonraki İKF'inin çıkartılması için 1. adımdan artıktık sinyali yeni giriş sinyali olarak alınarak tekrar işleme koyulur. Bu işlem artıktık sinyalinde zarfların elde edilmesine yetecek kadar uç noktası kalmayana kadar devam eder [15]. Artıklık (kalan) sinyali Denklem (3.16)'ya göre hesaplanır;

$$\text{kalan}(m,n)=\text{giriş}(m,n)-\text{içkin kip fonk}_i(m,n) \quad (3.16)$$

Yukarıda anlattığımız AKA adımlarına ait görsel Şekil 3.7'de yer almaktadır.



Şekil 3.7. Ampirik Kip Ayrışımına ait bir görsel [18]

3.2.3. İstatistiksel momentler

Girdi verilerini özellik kümesine dönüştüren bir boyut azaltma biçimi özellik çıkartma olarak adlandırılır. Çalışmada sınıflandırma yapabilmek için yüksek dereceli istatistiksel momentleri özellik olarak çıkartılmıştır. İstatistiksel momentler veri kümesi örneklerinin ortalamaya göre dağılımını, asimetrisini ve yoğunlaşmasını analize edebilmek için kullanılır. EEG verilerinin istatistiksel momentleri Denklem (3.17), (3.18) ve (3.19) formüllerine göre hesaplanmıştır;

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)^2, \mu = \frac{1}{N} \sum_{i=1}^N X_i \quad (3.17)$$

$$\beta_1 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)^3 / \sigma \quad (3.18)$$

$$\beta_2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)^4 / \sigma \quad (3.19)$$

N: noktalı X $\{X_1 X_2 X_3 \dots \dots \dots X_n\}$ veri için;

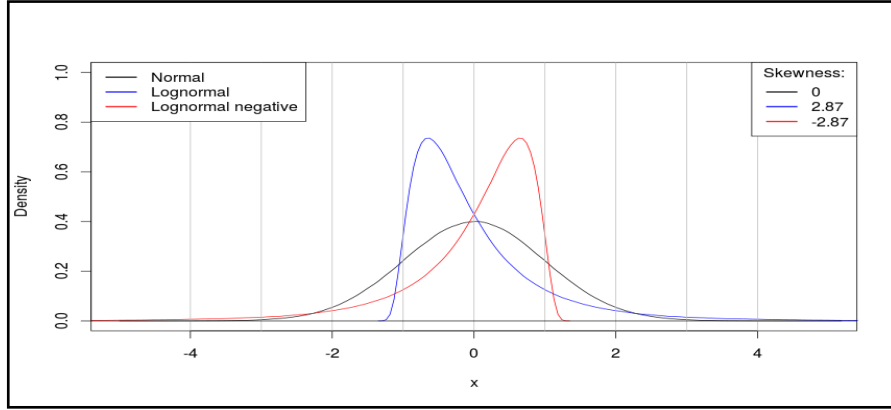
μ : ortalama

σ^2 : varyans

β_1 : çarpıklık

β_2 : basıklık değerleri

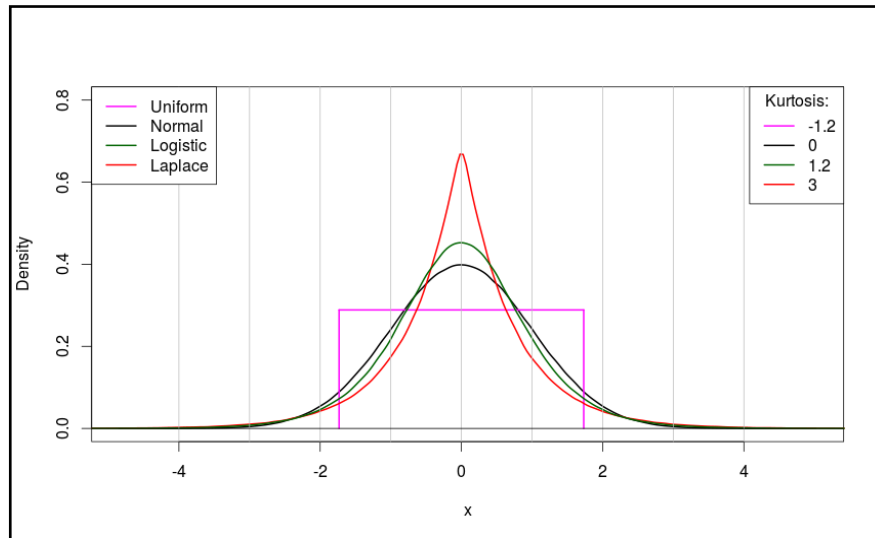
Ortalama ve varyans verinin birinci ve ikinci dereceli anlarını ifade ederken çarpıklık ve basıklık üçüncü ve dördüncü dereceli momentleri olarak değerlendirilebilir. Varyans başka bir ifade ile dağılım ölçüsüdür [8]. Veri setindeki değerlerin ortalamaya göre değişimini temsil etmektedir. Çarpıklık, örnek ortalamasının etrafındaki verilerin asimetrisinin ölçüsü diye ifade edilebilir. Şekil 3.8’de asimetri katsayıları ve dağılımları verilmiştir. Şekil 3.8 yorumlanacak olursa çarpıklık negatif olduğunda, verilerin ortalamasının soluna daha fazla yayıldığı görülmektedir, çarpıklık pozitifse, veriler sağa daha fazla yayılmış olarak görülür. Normal (simetrik) dağılımın çarpıklığı sıfırdır [25].



Şekil 3.8. Dağılımlar ve asimetri katsayıları [25]

Şekil 3.8’de rastgele değişkenlerin farklı standart dağılımları gösterilmektedir. Asimetrik katsayının değişimi hakkında dağılım kuyruğundaki uzunluğa bağlı olarak yorum yapılabilir. Uzunluğa bağlı olarak asimetrik katsayıyı değerlendirebiliriz [25].

Basıklık dağılımda bulunan aykırı değerlerin ölçüsüdür. Basıklık katsayısı aşağıdan sınırlı dolayısıyla değeri negatif olamaz. Genellikle normal dağılımın basıklıklığı 3’tür. Eğer basıklık katsayısı 3’ten büyükse, dağılımın normal dağılımdan daha kalın kuyrukları ve daha yüksek bir tepe noktası vardır. Eğer sıfırdan küçükse, daha ince kuyruklar ve daha düşük bir tepeye sahiptir [25]. Şekil 3.9’da basıklık katsayıları ve o katsayılara göre dağılımlar verilmiştir.



Şekil 3.9. Dağılımlar ve basıklık katsayıları [25]

Şekil 3.9’da farklı dağılımlar ve bunlara karşılık gelen basıklık değerleri gösterilmektedir. En kalın kuyruklu olanı Laplace dağılımıdır. Dağılımın zirvesinin yüksekliği katsayı değerinin büyüklüğü ile doğru orantılıdır [25].

3.3. Sınıflandırma

Bu çalışma da özellik çıkartımı yapılan veriye ait sınıflandırma işleminde MATLAB programından faydalanılmıştır. Verilerin bir kısmı eğitimde kullanılıp kurulan model aracılığı ile bir kısmı ise test amacıyla kullanılmıştır. Böylece oluşturulan modelin performansı değerlendirilmiştir.

Sınıflandırma işlemi için en çok kullanılan uygulamalardan biri denetimli makine öğrenmesidir. Bu uygulama hangi sınıf etiketi belli olan verilere ait özelliklerden yaptığı genelleştirme ile olayı çözümler ve hangi sınıfa ait olduğu belli olmayan veri hakkında yorum yaparak sınıf etiketi atama işlemi yapar. Bu sınıflandırma yöntemlerini şu şekilde sıralamak mümkündür;

- Karar Ağacı
- Diskriminant Analizi
- Destek Vektör Makineleri
- Lojistik Regresyon
- K-en Yakın Komşuluk
- Topluluk Sınıflaması

Çalışmada DVM, K-YK, KA algoritmaları kullanılarak sınıflama işlemi yapılmaya çalışılmıştır.

Etiketi belli olan verilere ait özellikler seçilip koordinat sistemi üzerinde dağılım görülebilir. Sınıflandırma işleminin sonuçlarını karışım matrisi temsil etmektedir. Bu da gerçek sınıflar ile tahmin edicinin temsil ettiği sınıfa ait matristir. Eğri altında kalan alan sınıflandırıcı gücünü göstermektedir.

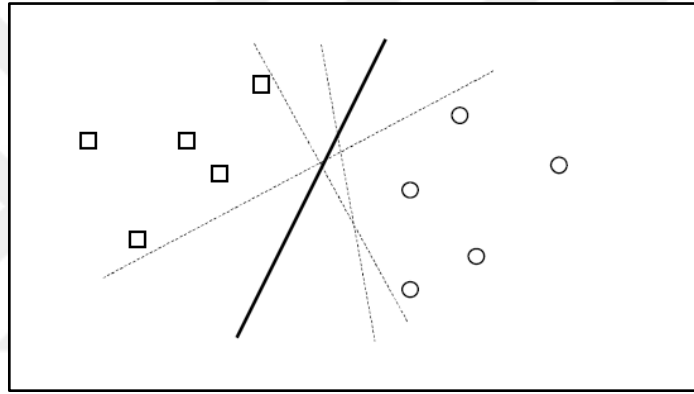
3.3.1. Destek vektör makineleri (DVM)

Gözetimli öğrenme algoritmasına dayanan makine öğrenmesi türüdür. İstatistiksel öğrenme tekniği kullanılmaktadır. İki ayrı sınıfa ait veriyi ayırt etme işleminde

kullanılır. Bu iki ayrı sınıfa ait veriyi ayırt etmek için en uygun olan düzlem belirleme işlemidir. Bu düzlem hiper düzlem olarak adlandırılmaktadır.

Veri setini doğrusal ya da doğrusal olmaması durumu DVM'inin ikiye ayrılmasında en önemli faktördür. Doğrusallık durumunda iki ayrı veri setinin hiper düzlemle ayırt etmek mümkündür fakat doğrusallığın olmadığı durumda veri setinin daha yüksek boyuttaki uzaya taşınmak gerekir. Uzaya taşınmasındaki amaç veriyi ayıracak çoklu düzlemin daha rahat bulunmasıdır [19].

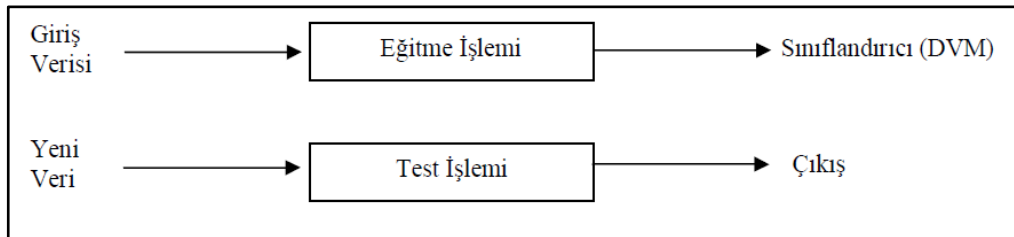
Optimum doğru kullanılarak sınıflandırma yapma amacıyla DVM kullanılır. Şekil 3.10'da optimum ayırıcı düzlemlerin görseli sunulmuştur.



Şekil 3.10. Optimum ayırıcı düzlemler [19]

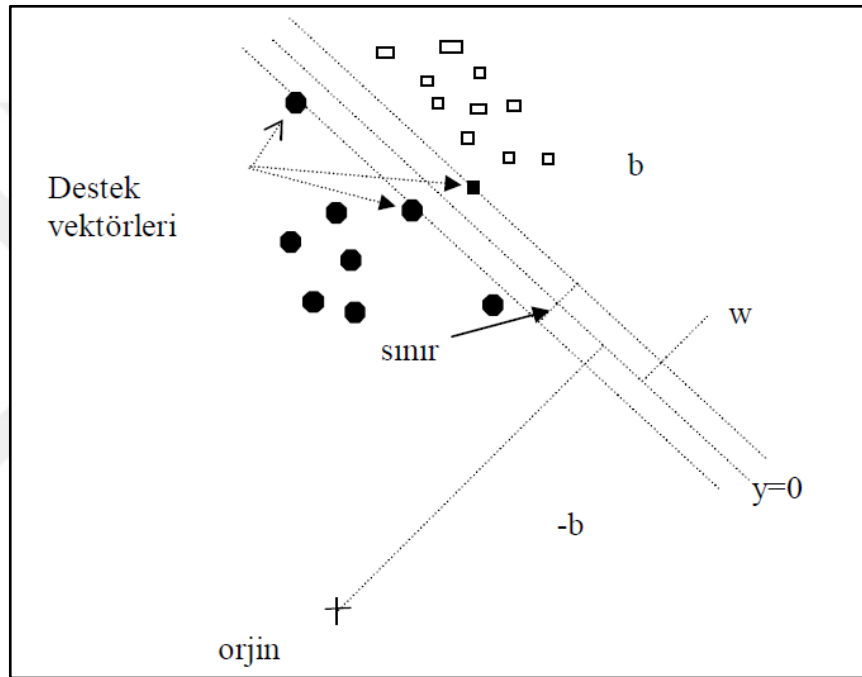
Yukarıda verilen Şekil 3.10'da iki sınıf arasında pek çok doğru çizmek mümkündür. Sınıflandırma da hata toleransını en düşük hale getirmek iki sınıfa da en uzak doğruyu bulmakla mümkündür.

DVM, uygun değer olan ayırıcı düzlemi bulmaya çalışırken önce optimizasyon problemini formülize etmeye çalışır. Daha sonra ikinci dereceli programlama denilen metotlar aracılığı ile bunu çözer [19].



Şekil 3.11. Makine öğrenmesi sınıflandırma mimarisi [19]

Şekil 3.11, bu mimari bütün sınıflandırmalara ait mimari olarak değerlendirilebilir. Eğitim ve test için ayrılan verilerin toplamı bize tüm veri seti kümesini vermektedir. Veri setinin eğitim için ayrılan kısmı oldukça önemlidir. Bu doğru sınıflandırma yapma işlemini doğrudan etkilemektedir. Bunlardan farklı veri setinin dağılımı da doğrudan etkilidir. Sınıflandırıcı model eğitim için ayrılan veriye dayanarak olayı öğrenir. Daha sonra sonuçları bilinen test verisiyle modelin çıktısı karşılaştırılır. Modelin doğruluğu ya da performansı bu iki değer arasındaki farka bağlı olarak değerlendirilir. Şekil 3.12’de DVM sınıflandırıcısı gösterilmektedir [19].



Şekil 3.12. DVM sınıflandırıcısı [19]

DVM’inde veri yapısına bağlı olarak iki durum söz konusudur. Bu verinin lineer veya lineer olmaması durumudur. Veriler arası ilişki lineer olduğunda bunu çözümlenmek oldukça kolaydır. Fakat lineer olmayan durumlar içinse bu veriler farklı bir uzaya taşınarak olaya çözüm bulunur.

3.3.1.1. Lineer ayrılma durumu

Eğitim için kullanılacak N elemandan oluşan verinin $X=\{x_i, y_i\}$, $i= 1,2,\dots,N$ olduğunu varsayalım. Burada $y_i \in \{1,-1\}$ etiket değerleri ve özellikler $X_i \in \mathbb{R}^d$ vektörüdür. Lineer olarak ayrıştırılan veriler, bu iki değerli veriler direkt olarak bir

aşırıdüzlem kullanılarak ayrılır. DVM’i aşırıdüzlemin değerlendirmeye alınan iki sınıfa ait örnek grubuna eş uzaklıkta olmasını amaçlamaktadır.

DVM doğrusal olarak ayrılmayan veriyi yüksek forma indirger, bu işlem aşırıdüzlemin bulunmasını kolaylaştırır. Bu işlemin yapılmasının amacı verilerin bir aşırıdüzlem ile ayrılmasını sağlamaktır. Aşırıdüzleme ait herhangi bir nokta, aşırıdüzlemin normali ve $b/\|w\|$ hiper uzayın orijine dik uzaklığı olmak üzere Denklem (20)’da gösterilen koşulu sağlar;

$$w^t \cdot x + b = 0 \quad (3.20)$$

DVM yönteminde araştırılan, örnekler lineer olarak ayrılabilirse, ve $y_i = 1$ etiketli örneklere eşit uzaklıkta olan optimum ayırıcı aşırıdüzlemin bulunmasıdır. Bunun için eğitim setinin aşağıdaki Denklem (3.21) ve (3.22) sağladığını varsayalım;

$$y_i = +1 \text{ için } w^t \cdot x + b \geq +1 \quad (3.21)$$

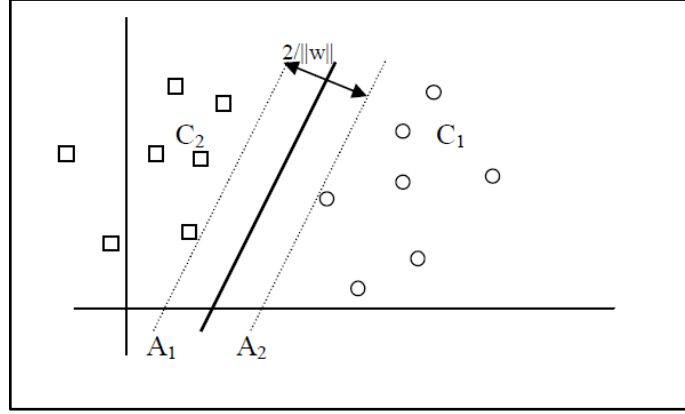
$$y_i = -1 \text{ için } w^t \cdot x + b \leq -1 \quad (3.22)$$

Bu eşitsizlikler bir arada değerlendirilecek olursa $i=1,2,\dots,N$ için Denklem (3.23) elde edilir;

$$y_i (w^t \cdot x + b) \geq +1 \quad (3.23)$$

Yukarıdaki koşula uyan aşırıdüzlemin iki tarafındaki en yakın örneklere olan dik uzaklıkları toplamı sınır olarak ifade edilmektedir. Optimum ayırıcı aşırıdüzlem, sınır maksimum yapan aşırıdüzlem anlamına gelmektedir ve bunun hesaplaması için b ve w gibi değerler hesaplanmalıdır. Şekil 3.13’te lineer ayrılabilir durumda optimum ayırıcı aşırıdüzlemin görseli verilmiştir.

Burada dikkat edilirse $A1$ ve $A2$ aşırıdüzlemleri eğitim verisi içermemektedir. Bu iki aşırıdüzlem arasındaki maksimum uzaklığın belirlenmesi için ise w değerinin minimize edilmesi gerekmektedir. DVM yöntemiyle amaç bu iki aşırıdüzlemin arasındaki uzaklığın (sınırın) maksimum olmasıdır. Aralarındaki sınır maksimum olan aşırıdüzlemlerin üzerinde yer alan noktalara ise destek vektörleri diye ifade edilmektedir.



Şekil 3.13. Lineer ayrılabilme durumunda optimum ayırıcı aşırıdüzlem [19]

Maksimum sınırın bulunması işlemi Denklem (3.24) ve (3.25) ile ifade edilir;

$$\min \frac{1}{2} \|w\|^2 \quad (3.24)$$

$$y_i (w^t x_i + b) \geq +1, \forall i \quad (3.25)$$

Denklem (3.24) ve (3.25) problemin çözümü sırasında kullanılacak formüllerdir. Başka bir ifade ile ikinci derece optimizasyon problemi denebilir. Problem çözümde ise Lagrange Formülasyonu kullanılır. Bunun yapılma sebebi ise formülasyona ait çarpanların kolay bulunması ve lineer olmayan durumlar için genelleme yapılabilmesidir [20].

Problemin Lagrange formülasyonu ise Denklem (3.26)'da verilmiştir;

$$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i y_i (w^t x_i + b) + \sum_{i=1}^N \alpha_i \quad (3.26)$$

Bu formülasyonda α_i değerleri pozitif Lagrange çarpanları olarak isimlendirilir. Probleme çözüm bulunması için Lagrange formülasyonunda yer alan ifade Karush-Kuhn-Tucker koşulları kullanılarak dual problemine dönüştürülmelidir.

3.3.1.2. Lineer ayrılamama durumu

Lineer ayrılmama durumunda sistem öğrenmek yerine ezberleme de yapabilir. Her türlü durumun değerlendirilmesi için C denen bir üst sınırın eklenmesi gerekmektedir. Bu Lagrange çarpanlarının alacağı maksimum değeri ifade

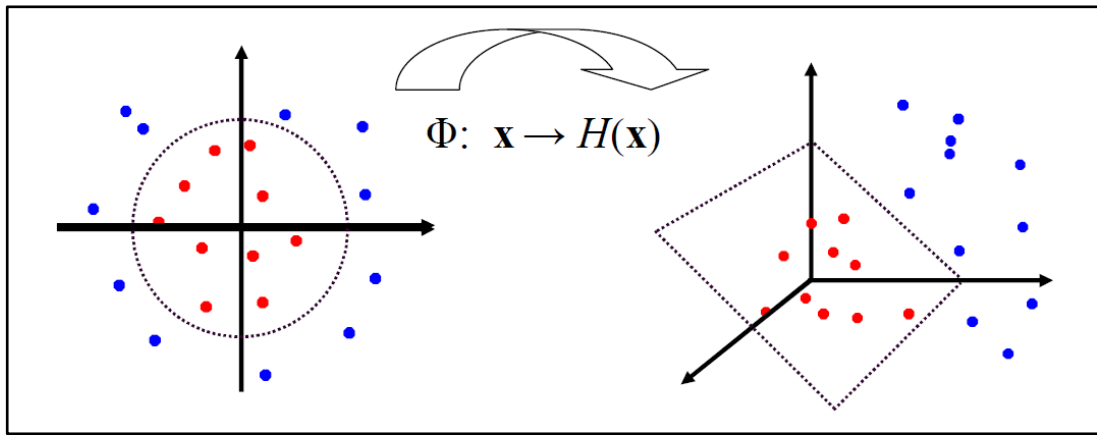
etmektedir. Bu yolla Lagrange çarpanları $0 \leq \alpha_i \leq C$ aralığında kalmaktadır. Bu düzenlemeye göre Lagrange Formülasyonu aşağıdaki Denklem (3.27)'de ifade edilir;

$$L_p = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum_i \alpha_i \{y_i(w^t x_i + b) - 1 + \xi_i\} - \sum_i \mu_i \xi_i \quad (3.27)$$

Yukarıdaki denklemde μ_i , ξ_i 'nin pozitif olmasını garanti etmek için kullanılmış olan Lagrange parametresidir. Probleme çözüm bulunması için Lagrange denklemde yer alan ifade yine aynı şekilde Karush-Kuhn-Tucker koşulları kullanılarak dual problemine dönüştürülmelidir [21].

3.3.1.3. Linear olmayan DVM

Yukarı da sınıflandırma problemlerinde örnekler lineer dağılıyorsa işlem kolaylığı olduğu lineer dağılmıyorsa da zayıflık parametresi ile problem çözüldüğü ifade edilmişti. Linear olmayan problemlerin çözümü için bir başka yol ise Çekirdek (Kernel) Fonksiyonlarıdır. Yani daha yüksek boyutlu ve lineer ayrılabilir bir uzaya taşınmasıdır. Çözüm yolu taşınma işleminin gerçekleştiği bu yeni uzayda aranır. Şekil 3.14'te çok boyutlu uzaya taşınma görseli sunulmaktadır [22].



Şekil 3.14. Giriş uzayının çok boyutlu uzaya dönüşümü [22]

3.3.2. K-en yakın komşuluk (K-YK)

K-YK hem regresyon hem de sınıflandırma problemleri için kullanılan gözetimli öğrenme algoritmalarından biridir. 1967 yılında T. M. Cover ve P. E. Hart tarafından ortaya konulmuştur. Sınıfları belli olan veriye dayanarak öğrenme olayı gerçekleşir. Veriye eklenecek yeni verinin mevcut veriye uzaklığı hesaplanarak en yakın

komşular seçilir. Uzaklık hesaplama işlemi için de sırasıyla Euclidean, Manhattan, Minkowski Uzaklık fonksiyonlarına göre yapılmaktadır.

Basit bir öğrenme algoritması taşıyan sınıflandırma yöntemidir. Test için ayrılan veri var olan eğitim kümesine eklenmektedir. Bütün verinin en yakın k adet komşusu dikkate alınarak bu komşular arasında en çok örneği bulunan sınıf test sınıfı olarak değerlendirmeye alınmaktadır. Çok büyük veri setleri için ideal bir yöntem değildir. Eğitim kümesine eklenen test verisinin sınıfı belirlenirken dikkate alınacak komşu sayısı K olarak ifade edilmektedir [26]. Çalışma mantığına dair adımlar görsel olarak aşağıdaki Şekil 3.15'te verilmiştir.



Şekil 3.15. K-YK adımları [26]

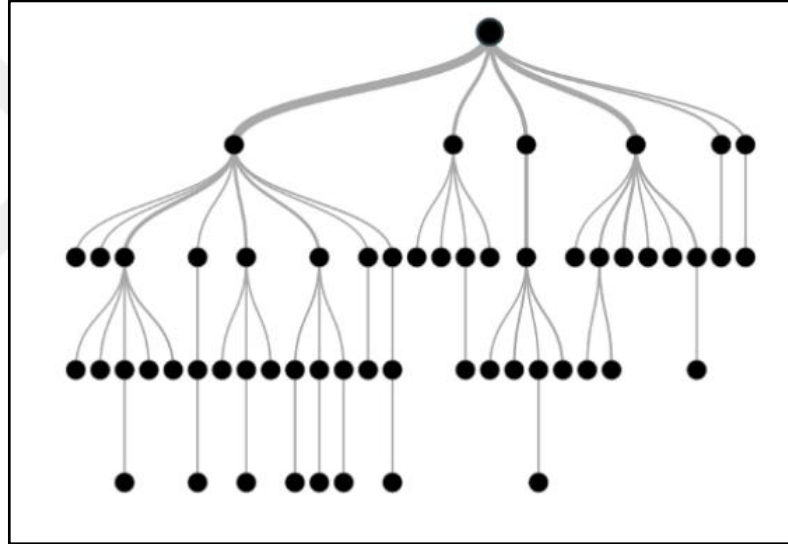
Şekil 3.15 yorumlanacak olursa çalışma prensibi adımları şöyle sıralanabilir;

- K parametresi dikkate alınan noktaya en yakın olan nokta sayısını ifade etmektedir.
- Uzaklık fonksiyonları kullanılarak veri setine ilave edilecek verinin var olan veriye uzaklığı tek tek hesaplanır.
- Bu uzaklıklar dikkate alınarak en yakın K komşu alınır. Çıkarılan özniteliklere göre uzaklıklardan en yakın K komşu ele alınır ve sınıf atama işlemi gerçekleştirilir.
- Belirlenen sınıf tahmini istenen veriye ait gözlem değerinin sınıfını olarak belirenmiş olur. Başka bir deyişe yeni veriye etiket konulmuş olur [26].

3.3.3. Karar Ağacı (KA)

1970'li yıllarda ortaya konan ilk algoritma ve yazılım olan KA algoritması Michigan Üniversitesi'nden Morgon ve Sonquist tarafından yazılmıştır. Zamanla bu algoritma

türleri geliştirilmiştir ve en çok kullanılanlara ID3, C4.5, CHAİD, CART örnek verilebilir. Ağaç yapısı şeklinde olduğundan dolayı isimlendirmesi de bu şekilde oldurmuştur. Çalışma prensibinde tümevarım yaklaşımı mevcuttur. Basit olmasına karşın çok yaygın bir şekilde kullanımı hala da devam etmektedir. Çok tercih edilmesinin temel sebeplerinden biri de gürültüye karşı oldukça ayırt edici olmasıdır. KA elemanları sırasıyla iç düğüm elemanları, dal ve uç yapraklar olarak ayrılabilir. Test işleminin yapıldığı ve sorular vasıtasıyla ne tarafa yöneleceğinin belirlendiği düğümlerdir. Sorulan bu sorulara verilen cevap ise dal olarak isimlendirilmektedir. Sınıf etiketlerinin hangi kategoriye ait olduğu hakkında bilgi almamızı sağlamakta olan kısım ise uç yaprak olarak nitelendirilir. Yukarıdan aşağıya doğru inen bir çalışma mantığı mevcuttur [27]. Bir KA örneği aşağıda Şekil 3.16'da verilmiştir.



Şekil 3.16. Karar Ağacı örneği [27]

Günümüzde de özellikle veri madenciliğinde kullanımı devam etmektedir. Temel olarak düğüm, dal ve yaprak olmak üzere üç kısımdan oluşan makine öğrenmesinde kullanılan sınıflandırma yöntemlerinden biridir. Düğümler çıkarılan öznitelikleri temsil etmektedir. En yukarıdaki düğüm kök düğüm diye ifade edilen verileri gösteren düğüm olmaktadır. Yaprak kısmı ise sınıf etiketlerini göstermek amacıyla kullanılmaktadır.

Olayı en basit şekilde çözümlenecek karar verme adımları ile çok fazla kayıt gerektiren verileri daha az kayıtle sınıflandırma işlemleri için ideal bir yöntem olmaktadır. Bu yöntem sınıflandırma ve regresyonda kullanılmaktadır. Doğruluğu

etkileyen şey bölünme kriteridir ve bu kriterler regresyon ve sınıflandırma için farklılık arz etmektedir. Yukarıdaki KA örneği incelendiğinde alt düğümlerin oldukça fazla olduğu görülür bunun temel sebebi ise işleme dair homojenliği arttırmaktır. Kategorik değişkenler için farklı algoritmalar mevcuttur bunlara örnek entropi, gini ve en küçük kareler yöntemleri verilebilir. Belirsizlik ölçütünü biz entropi olarak tanımlamaktayız. Amaç entropiyi en aza indirgeyecek şekilde veriyi bölebilmektir. Bölmeler ne kadar fazla ise tahmin o kadar güçlü olmaktadır.

Claude Shannon tarafından öne sürülen entropi Denklemi (3.28)'de verilmektedir;

$$H = - \sum P(X) \cdot \log P(X) \quad (3.28)$$

$P(X)$: belirli bir sınıfa ait grubun yüzdesini ve

H : entropi değerini göstermektedir.

Doğru bölünmelerin doğru tahminde etkili olduğu düşünülürse bu bölünmeler sonuç açısından en önemli faktör haline almaktadır. Bu bölünme için ise bilgi kazancı kullanılmaktadır. Bilgi kazancı ise Denklem (3.29) ile verilen eşitlik baz alınarak hesaplanmaktadır;

$$\text{kazanç}(S, D) = H(S) - \sum_{V \in D} (|V|/|S|) \cdot H(V) \quad (3.29)$$

S : orijinal veri,

D : kümenin bölünmüş parçası,

V : ise S 'yi oluşturan ayrık küme

Bu durumda kazanç orijinal veri setine ait entropi ile özniteliklere ait entropi değerleri arasındaki fark olarak ifade edilebilir. Sınıflandırma problemlerinde veri seti eğitim ve test olarak ayrıştırılır. Model eğitim verisine dayanarak oluşturulur. Bu modelin tutarlılığı test verisi ile sağlanarak probleme çözüm bulunur [27].

4. DENEYSEL SONUÇLAR

4.1. EEG Verilerinden Özellik Çıkartımı

Bu tez kapsamında EEG verilerinden özellik çıkartımı gerçekleştirebilmek için öncelikle frekans uzayına geçilmektedir. Bu amaçla da iki yaklaşım kullanılmaktadır. Çalışmada kullanılan birinci yaklaşım ADD kullanılarak frekans uzayına geçmektedir. Bu yöntem EEG işaretlerini alt bantlara ayırmaktadır. Yapılan literatür taramasında 4. dereceden Daubechies dalgacık fonksiyonu epileptik EEG sinyal analizi için en uygun seviye derecesi olarak seçildiğinden [23], bu çalışmada da aynı yol izlenmiştir. Örnekleme frekansı 173,6 Hz olan 4 seviyeli ADD ayrışmasına cevap veren frekans bantları, Tablo 4.1’de gösterilmiştir. Şekil 4.1’de ise nöbet anında bir EEG sinyalinden çıkarılan alt bantlar gösterilmiştir.

Tablo 4.1. EEG sinyallerinin 4 dereceli ADD ile ayrıştırıldığında elde edilen detay (D) ve yaklaşık (A) katsayıları [23]

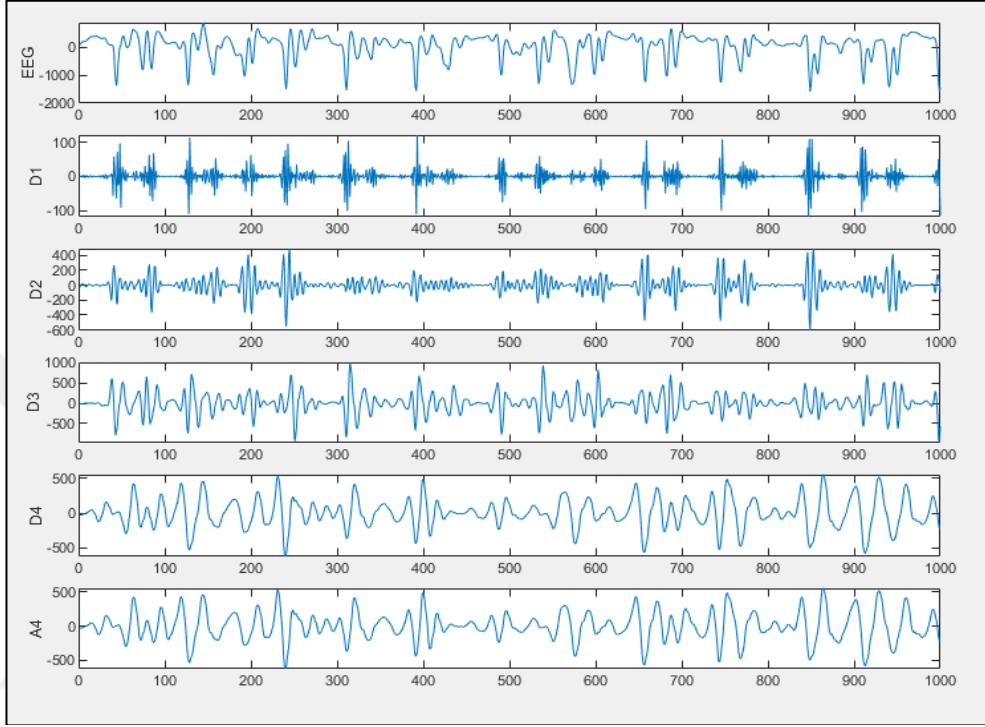
Ayrıştırma seviyesi	Katsayı	Bant ismi	Frekans aralığı (Hz)
1	D1	Gama	43,4-86,8
2	D2	Beta	21,7-43,4
3	D3	Alfa	10,8-21,7
4	D4	Teta	5,4-10,8
4	A4	Delta	0-5,4

Bu tezde frekans uzayına geçmek için kullanılan diğer yöntem ise AKA’dır. Bu yöntem uygulanıp EEG verisini 5 tane İKF’e ayırmaktadır. 1. İKF en yüksek frekansa sahipken kalan sinyal da en düşük frekansı içermektedir. Tablo 4.2’de ilk 6 İKF’lerin kapsadığı frekans aralığı verilmektedir.

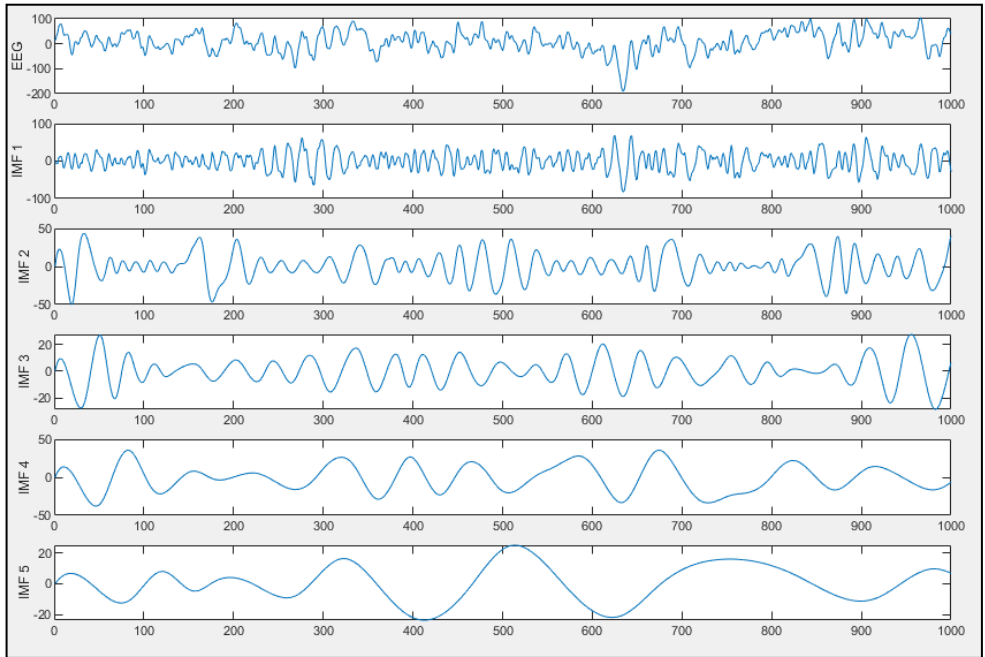
Tablo 4.2. İKF’lerin kapsadığı frekans aralığı [15]

İçkip Kip Fonksiyonu	Frekans aralığı (Hz)
1	0-45
2	0-30
3	0-20
4	0-10
5	0-7
6	0-3

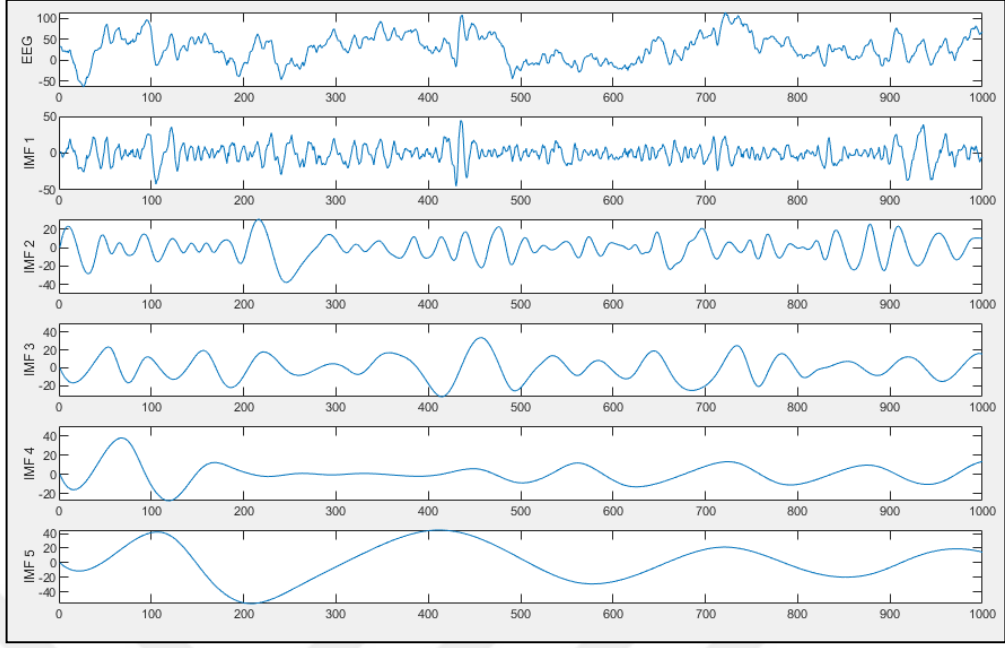
Nöbet anı EEG'sinin temel özellikleri ilk beş İKF'te yer almaktadır [15]. Dolayısıyla her EEG segmentinin İKF1- İKF5'leri EEG özelliklerini çıkarmak için kullanılır. Şekil 4.2, Şekil 4.3, Şekil 4.4'te sırasıyla sağlıklı, nöbet öncesi ve nöbet anına ait EEG sinyallerinin ilk beş İKF'leri verilmektedir.



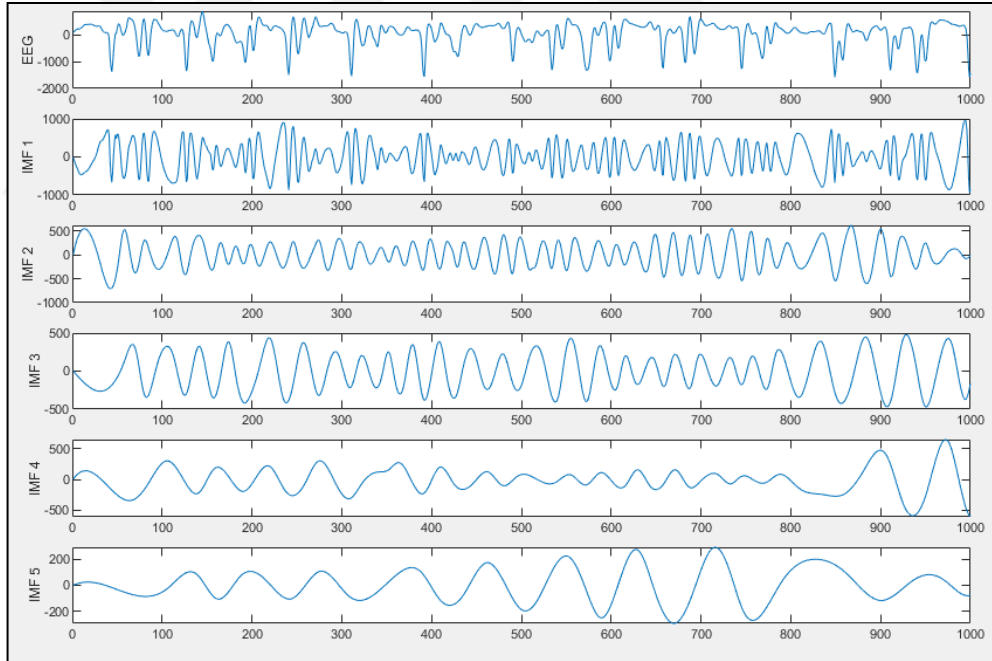
Şekil 4.1. Nöbet anındaki EEG sinyalinden ADD yöntemiyle elde edilen alt bantlar



Şekil 4.2. A kümesine ait sağlıklı EEG döneminin ilk beş İKF'i



Şekil 4.3. D kümesine ait nöbet öncesi EEG döneminin ilk beş İKF'i



Şekil 4.4. E kümesine ait nöbet anı EEG döneminin ilk beş İKF'i

Son olarak, ADD'dan çıkartılan detay ve yaklaşık katsayıları ve AKA'dan çıkarılan İKF'ler kullanılarak her biri için istatistiksel moment özellikleri belirlenmiştir.

Nöbet anına (E kümesine) ait EEG verisinin AKA yöntemi ile hesaplanan istatistiksel momentleri Tablo 4.3'te verilmiştir.

Tablo 4.3. AKA yöntemi ile hesaplanan nöbet anına ait istatistiksel momenter

Veri seti	Özellikler	IKF 1	IKF 2	IKF 3	IKF 4	IKF 5
E seti	Ortalama	33,084	65,251	-0,861	-7,373	-7,833
	Varyans	98365,592	83427,738	32050,131	11252,06	4916,921
	Çarpıklık	0,0193	-0,125	-0,046	0,029	0,204
	Basıklık	3,512	2,390	2,982	3,263	2,224

4.2. EEG İşaretlerinde Epilepsi EEG Tespiti ve Tahmini

Deneyde kullanılan Bonn Üniversitesi veri tabanından alınan EEG verileri (A, B, C, D, E kümeleri) ile farklı senaryolar oluşturularak epilepsi EEG tespiti ve tahmini yapılmıştır. Özellik çıkartımı kısmında elde edilen ortalama, varyans, çarpıklık ve basıklık olmak üzere çıkarılan özellikler DVM, K-YK, KA yöntemlerine göre sınıflandırılmıştır. Senaryolara ait açıklama Tablo 4.4'te verilmektedir.

Tablo 4.4. Farklı senaryoların açıklanması

Senaryolar	Kümelerin gruplandırılması	Sınıflandırma
Durum I	A, B kümeleri C,D kümeleri E kümesi	Sağlıklı Nöbet öncesi Nöbet anı
Durum II	A kümesi D kümesi E kümesi	Sağlıklı Nöbet öncesi Nöbet anı
Durum III	A kümesi E kümesi	Sağlıklı Nöbet anı
Durum IV	ABCD kümeleri E kümesi	Sağlıklı Nöbet anı
Durum V	D kümesi E kümesi	Nöbet öncesi Nöbet anı

Sınıflandırma yapılırken veri kümesinin %60'ı eğitim için ve %40 test için ayrıştırılmıştır. Tablo 4.5'te ADD yöntemi baz alınarak elde edilen doğruluk değerleri verilmiştir. Aynı şekilde Tablo 4.6'da AKA ile elde edilen doğruluk değerleri verilmiştir.

Tablo 4.5. ADD yöntemine göre elde edilen doğruluk değerleri

Senaryolar/ Sınıflandırıcı	DVM	K-YK	KA
Durum I	40	60	62,5
Durum II	93,33	95	93,33
Durum III	100	97,5	100
Durum IV	95,5	97	95,5
Durum V	96,25	95	98,75

Tablo 4.6. AKA yöntemine göre elde edilen doğruluk değerleri

Senaryolar/ Sınıflandırıcı	DVM	K-YK	KA
Durum I	60	59,5	48,5
Durum II	86,67	95	96,67
Durum III	100	100	96,25
Durum IV	94,5	95,5	99
Durum V	96,25	93,75	97,5

Sınıflandırma yapmak için ikinci yöntem olarak MATLAB programının Classification Learner uygulaması kullanılmıştır. Tablo 4.7’de ADD yöntemi baz alınarak elde edilen doğruluk değerleri verilmiştir. Aynı şekilde Tablo 4.8’de AKA ile elde edilen doğruluk değerleri verilmiştir.

Tablo 4.7. ADD yöntemine göre tüm bantlar için elde edilen doğruluk değerleri

Senaryolar/ Sınıflandırıcı	DVM	K-YK	KA
Durum I	68,6	68	82,4
Durum II	90,7	82	91,7
Durum III	98	96,5	99,5
Durum IV	94,6	95,6	96,4
Durum V	92,5	93	98

Tablo 4.8. AKA yöntemine göre tüm İKF’ler için elde edilen doğruluk değerleri

Senaryolar/ Sınıflandırıcı	DVM	K-YK	KA
Durum I	78	69,8	79,6
Durum II	89,7	80,3	93,7
Durum III	94	83,5	99,5
Durum IV	95,4	93,2	98,8
Durum V	91,5	86,5	96

4.3. EEG İşaretlerine Dair Analiz Sonuçları

Sınıflandırma işleminin sonuçlarını karışım matrisi temsil etmektedir. Bu da gerçek sınıflar ile tahmin edicinin temsil ettiği sınıfa ait matristir ve Tablo 4.9’da bu matris verilmektedir.

Tablo 4.9. Karışım matrisi [28]

	Tahmin edilen Pozitif	Tahmin edilen Negatif
Pozitif	TP	FN
Negatif	FP	TN

Tablo 4.9’da yer alan TP aslında pozitif olan ve model tarafından da pozitif olarak tahmin edilen değerlerdir. FN ise aslında pozitif olan, ancak model tarafından da

negatif olduğu tahmin edilen değerlerdir. FP, aslında negatif olan, ancak model tarafından pozitif olarak tahmin edilen değerlerdir. TN de aslında negatif olan ve model tarafından da negatif olduğu tahmin edilen değerlerdir [28].

Kesinlik, temel olarak model tarafından pozitif olarak sınıflandırılan sonuçlardan, gerçekte kaçının olumlu olduğunu göstermektedir. Denklem (4.1)'e göre kesinlik değeri hesaplanır;

$$\text{Kesinlik (Precision)} = \frac{TP}{(TP+FP)} \quad (4.1)$$

Geri çağırma model tarafından kaç tane gerçek pozitif değer (pozitif olarak işaretlenmiş noktaların) hatırlandığını veya bulunduğunu ifade etmektedir. Denklem (4.2)'de geri çağırma formülü gösterilmektedir;

$$\text{Geri çağırma (Recall)} = \frac{TP}{(TP+FN)} \quad (4.2)$$

F1-score başka bir değişle doğruluk değeri, kesinlik ve geri çağırmanın harmonik ortalamasıdır. Denklem (4.3)'te hassasiyet ve hatırlama arasındaki dengeyi ifade etmektedir;

$$\text{Doğruluk (F1 score)} = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (4.3)$$

ROC eğrisi veya alıcı işletim karakteristiği, tüm bilgileri özetlemek için basit bir yol sağlamakta olan grafiklerdir. Bu grafiğin anlamı da her bir eşik ürettiği tüm karışım matrisilerini özetlemektedir. ROC eğrisi, karar verme aşamasında en iyi eşik belirlemektedir.

AUC veya eğri altındaki alan, hangi yönteminin daha iyi olduğuna karar vermeye yardımcı olur. ROC eğrisi altında kalan alan sınıflandırıcı gücünü göstermektedir.

Analizler sonucunda en iyi performansı sergileyen modellere ait karışım matrisi ve ROC eğrisine ait grafikler aşağıda verilmiştir. ADD yöntemi sayesinde farklı durumlarda elde edilen en iyi model ve o modele ait karışım matrisi ve ROC, AUC eğrileri Şekil 4.5-4.19'da gösterilmektedir. Şekil 4.5- 4.6'da I.durum için elde edilen KA modeline ait karışım matrisi verilmiştir. Seçilen sınıflandırıcının her bir sınıfta nasıl bir performans gösterdiğini anlamak için karışım matrisi grafiği Classification

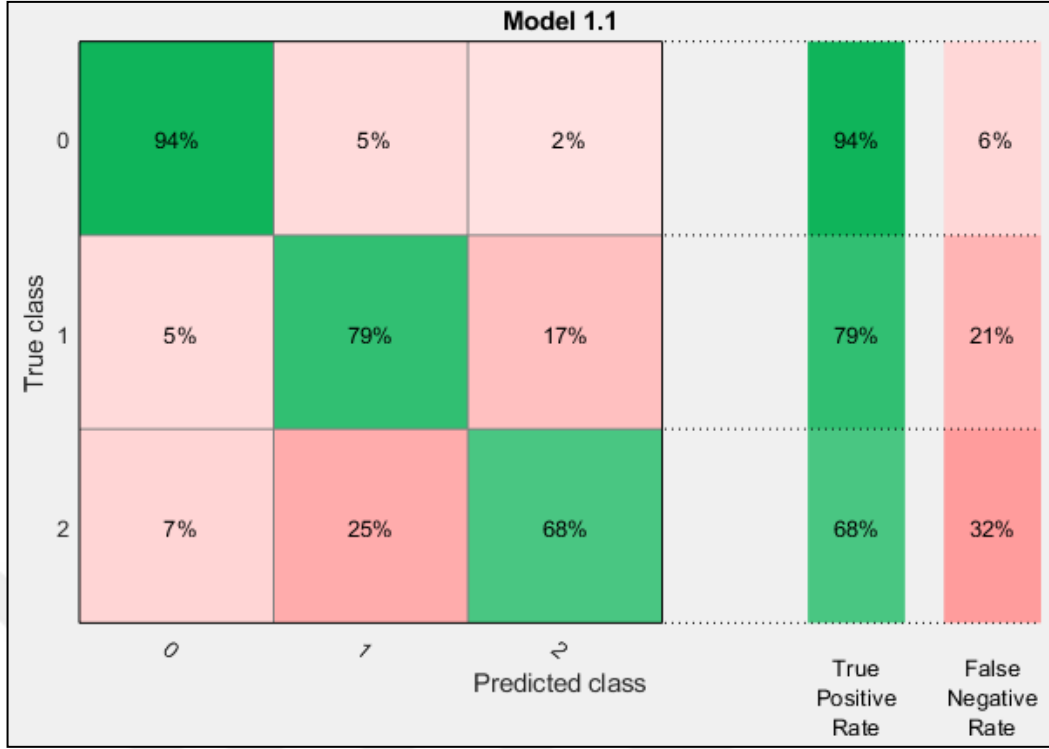
Learner ile elde edilmiştir. Bu matrisle bir modelin eğitim sonrasında her bir sınıf için başarımlarını değerlendirilmektedir [29].

		Model 1.1		
		0	1	2
True class	0	187	9	4
	1	10	157	33
	2	7	25	68
		0	1	2
		Predicted class		

Şekil 4.5. Gözlenen örnek sayısına göre durum I için KA modeline ait karışım matrisi

Şekil 4.5'te her hücrede gözlemlenen örnek sayısı gösterilmiştir. Burada satırlar gerçek sınıfı gösterir, sütunlar ise tahmin edilen sınıfı göstermektedir. Çapraz hücreler gerçek sınıf ile tahmin edilen sınıfın eşleştiği yeri göstermektedir. Eğer bu çapraz hücreler yeşilse, sınıflandırıcı iyi bir performans göstermiştir ve gerçek sınıfın gözlemlenen örnekleri doğru bir şekilde sınıflandırılmış olduğunu ifade eder [29].

1. satırda KA sınıflandırıcısı sağlıklı sınıfına ait 200 örnekte oluşan EEG verinin 187'sini sağlıklı olarak doğru sınıflandırmıştır. Kalan 9'nu nöbet öncesi ve 4'ünü de nöbet anı olarak yanlış sınıflandırmıştır. Aynı şekilde yorumlayacak olursak, model 2. satır da nöbet öncesi sınıfına ait 200 örnek içeren EEG kaydınının 157'si nöbet öncesi olarak doğru sınıflandırmıştır, 10'unu sağlıklı ve 33'ünü nöbet anı olarak yanlış sınıflandırmıştır. Son satırdaki 100 örnekte oluşan nöbet anına ait EEG'nin 68'ini model doğru tahmin etmiştir. Model diğer kalan 7'sini sağlıklı ve 25'ini nöbet öncesi olarak yanlış tahmin etmiştir.

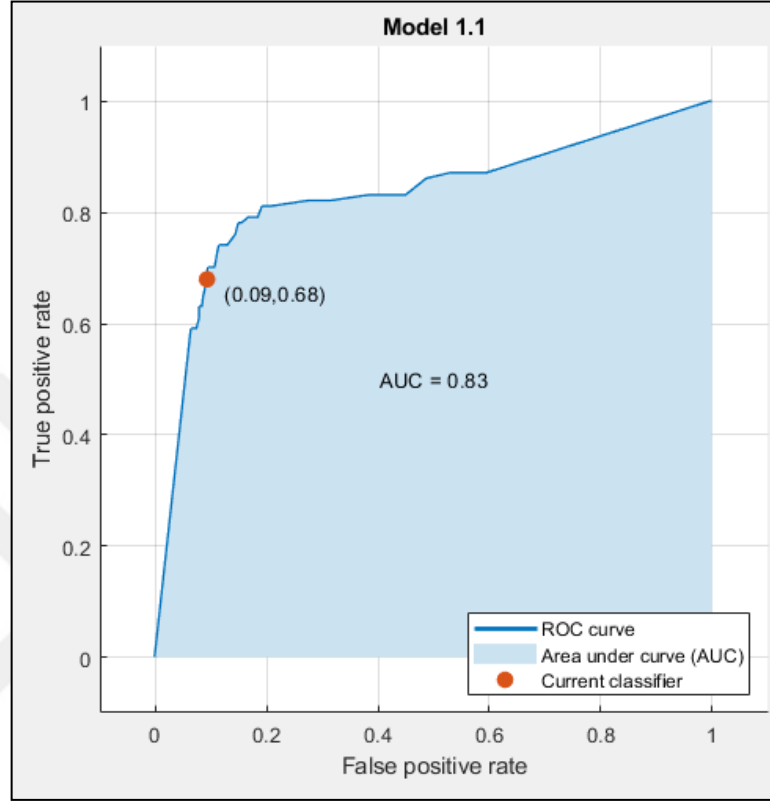


Şekil 4.6. Durum I için KA modeline ait karışım matrisi

Şekil 4.6’da I.durum için elde edilen KA modeline ait karışım matrisi verilmiştir. Bu örnekte, Bonn veri setini kullanarak, en üstteki satır gerçek sağlıklı sınıfı olan tüm EEG kayıtlarını gösterir. Sütunlar ise tahmin edilen sınıfları gösterir. Şeklin en sağındaki iki sütun Doğru Pozitif Oranlar, Yanlış Negatif Oranlarıdır ve bu değerler sınıflandırıcının sınıf başına nasıl bir performans elde ettiğini göstermektedir. Sağlıklı EEG kayıtlarındaki bazı örnekleri model yanlış sınıflandırmıştır. Yani %5’ini yanlış olarak nöbet öncesi ve %2’sini yanlış bir şekilde nöbet anı olarak sınıflandırmıştır. En son sütündeki açık pembe ile gösterilen %6 oranı, bu sınıftaki yanlış sınıflandırılmış Yanlış Sınıf Oranını ifade etmektedir. Nöbet öncesi için EEG’lerin %5’i sağlıklı ve %17’si nöbet anı olarak yanlış sınıflandırılmıştır. Yanlış Sınıf Oranını da %21 olarak verilmiştir. Nöbet anı EEG’lerinin %7’si sağlıklı ve %25’i nöbet öncesi olarak yanlış sınıflandırılmıştır. Yanlış Sınıf Oranını da %32 olarak verilmiştir.

En üst sırada, sağlıklı EEG kayıtlarının %94’i doğru bir şekilde sınıflandırıldığını göstermekte ve bu sınıf için Gerçek Pozitif Oranını ifade etmektedir. Aynı şekilde 2. satır nöbet öncesi sınıfı olan EEG kayıtlarıdır ve model %79’unu doğru bir şekilde sınıflandırdığını göstermektedir. En son satırda nöbet anı sınıfı olan EEG kayıtları

yer almaktadır. Model Gerçek Pozitif Oranı %68 olarak göstermektedir. Bu 3'lü sınıflandırma problemini çözmekte KA modelinin I. durum için doğruluğu %82,4 olarak hesaplamıştır.



Şekil 4.7. Nöbet anı pozitif sınıfı olduğunda durum I için KA sınıflandırıcısına ait ROC ve AUC eğrileri

Durum I için KA sınıflandırıcısına ait ROC ve AUC eğrileri olan Şekil 4.7'nin üzerindeki işaretleyici seçili olan sınıflandırıcının performansını göstermektedir.

İşaretçi, seçili olan sınıflandırıcı için hatalı pozitif oranı (FPR) ve gerçek pozitif oranı (TPR) değerleridir. Şekildeki 0,09'luk bir hatalı pozitif oran, mevcut sınıflandırıcının gözlemlenen örneklerin %9'unu hatalı olarak pozitif sınıfa atadığını gösterir. 0,68'lik gerçek pozitif oran, sınıflandırıcının gözlemlenen örneklerin %68'ini pozitif sınıfa doğru şekilde atadığını göstermektedir.

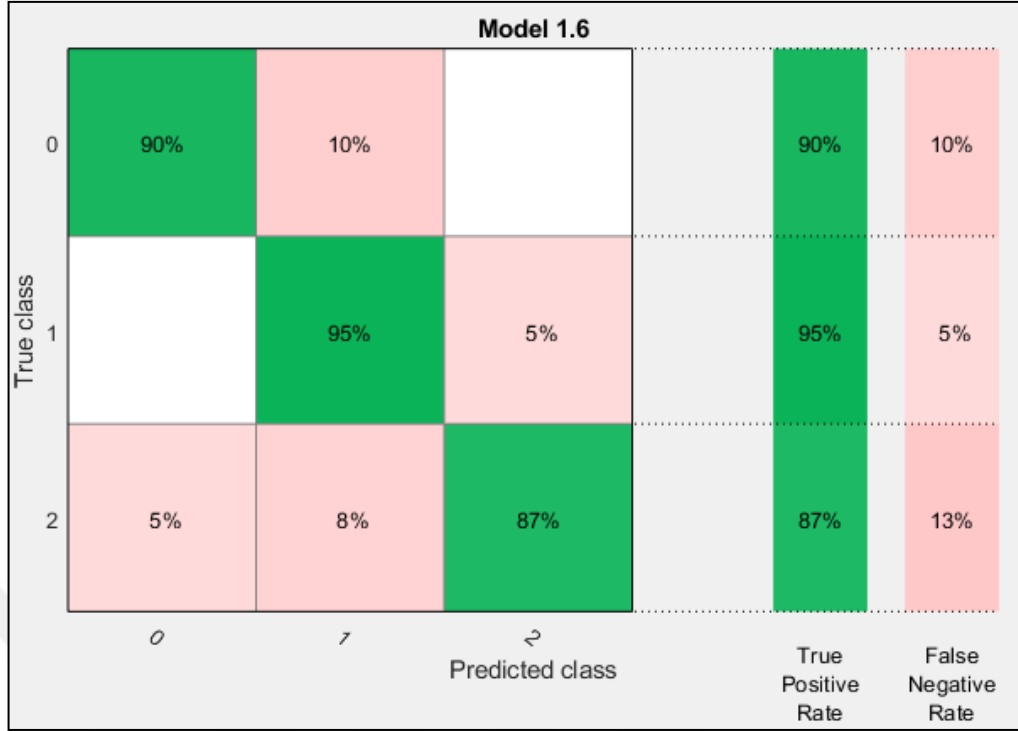
Şekilde AUC 0,83 değerini göstermektedir. Eğri Altındaki Alan sınıflandırıcının kalitesinin bir ölçüsüdür. Şekildeki eğri altındaki alan ne kadar büyük ise seçilen modelin iyi bir performansa sahip sınıflandırıcı olduğunu gösterir.

		Model 1.6		
		0	1	2
True class	0	90	10	
	1		95	5
	2	5	8	87
		0	1	2
		Predicted class		

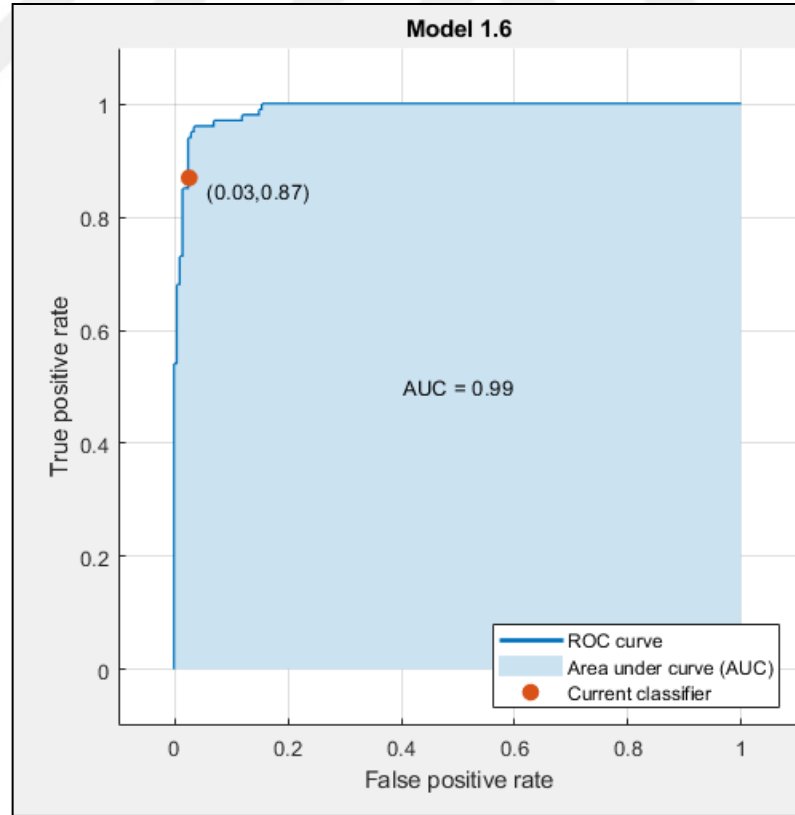
Şekil 4.8. Gözlenen örnek sayısına göre DVM sınıflandırıcısına ait II. durum için karışım matrisi

Şekil 4.8'de 1.satırda II. durum için DVM sınıflandırıcısı sağlıklı sınıfına ait 100 örnekten 90'nı sağlıklı olarak doğru sınıflandırmıştır ve kalan 10'unu nöbet öncesi olarak yanlış sınıflandırmıştır. Aynı şekilde sınıflandırıcı 2. satır da nöbet öncesi sınıfına ait 100 örneğin 95'ni nöbet öncesi olarak doğru sınıflandırmıştır, kalan 5'ini ise nöbet anı olarak yanlış sınıflandırmıştır. Sınıflandırıcı 100 örnekten oluşan nöbet anı sınıfının 87'sini doğru tahmin etmiştir. Kalan 5'ni sağlıklı ve 8'ini nöbet öncesi olarak yanlış tahmin yapmıştır. Bu çapraz hücreler yeşil demek ki sınıflandırıcı iyi bir performans göstermiştir ve gerçek sınıfın gözlemlenen örnekleri doğru bir şekilde sınıflandırmış olduğunu ifade etmektedir.

Şekil 4.9 en üstteki satırda II. durum için sağlıklı EEG verilerinin % 90'ı sınıflandırıcı doğru tahmin ettiğini göstermekte ve bu sınıf için Gerçek Pozitif Oranını ifade etmektedir. Nöbet öncesi sınıfı olan EEG kayıtları 2. satırdadır ve model % 95'ini doğru tahmin ettiğini göstermektedir. Nöbet anı sınıfı olan EEG kayıtları da 3. satırda yer almaktadır. DVM sınıflandırıcı bu sınıf için Gerçek Pozitif Oranını %87'i olarak göstermektedir. Bu 3'lü sınıflandırma problemini çözmekte DVM sınıflandırıcısı II. durum için doğruluk değerini % 90,7 olarak elde etmiştir.



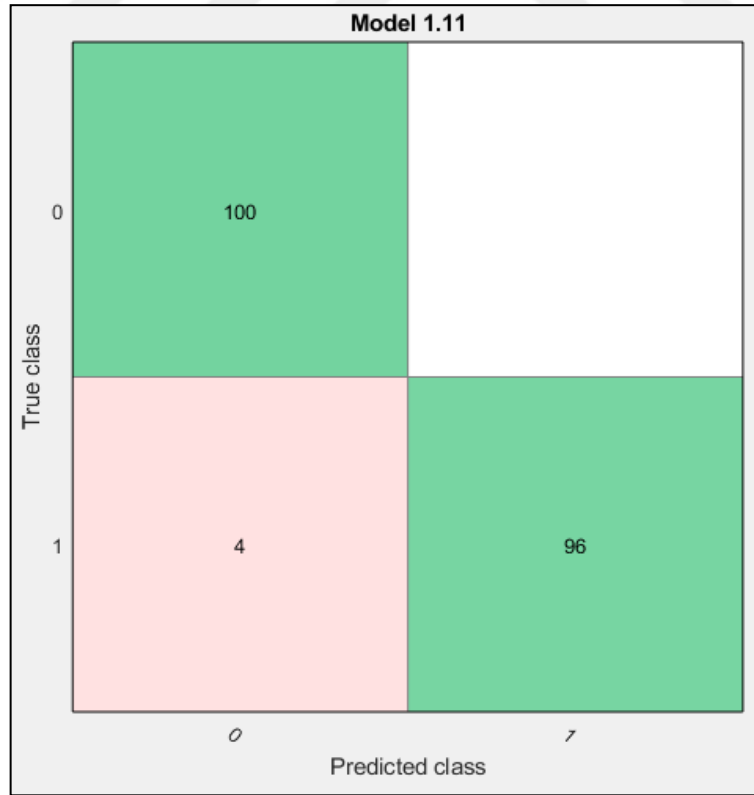
Şekil 4.9. Durum I için DVM sınıflandırıcısına ait karışım matrisi



Şekil 4.10. Nöbet anı pozitif sınıfı olduğunda durum II için DVM modeline ait ROC ve AUC eğrileri

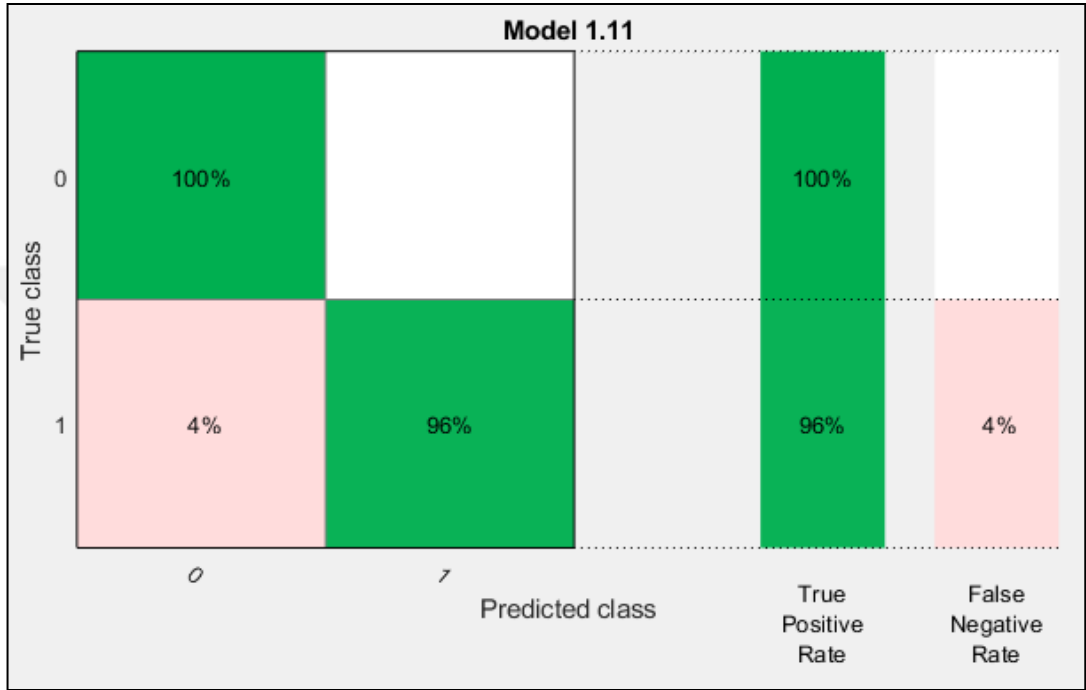
Şekil 4.10'da II. durum için DVM sınıflandırıcısına ait ROC ve AUC eğrilerinde olan turuncu işaret seçili olan sınıflandırıcı için hatalı pozitif oranı (FPR) ve gerçek pozitif oranı (TPR) değerlerini göstermektedir. Şekildeki 0,03'lük bir hatalı pozitif oranı, mevcut sınıflandırıcının gözlemlenen örneklerin %3'ünü hatalı olarak pozitif sınıfa atadığını gösterir. 0,87'lik gerçek pozitif oranı da sınıflandırıcının gözlemlenen örneklerin %87'sini pozitif sınıfa doğru bir şekilde atadığını ifade etmektedir. AUC 0,99 değerini almakta ve eğri altındaki alan büyük olduğundan dolayı seçilen modelin iyi bir performansa sahip sınıflandırıcı olduğunu göstermektedir.

Şekil 4.11'de III. durum için 1.satırda DVM sınıflandırıcısı sağlıklı sınıfa ait 100 örnekten tümünü sağlıklı olarak doğru sınıflandırmıştır. Aynı şekilde sınıflandırıcı 2. satırda nöbet anı sınıfa ait 100 örneğin 96'sını doğru sınıflandırmıştır, kalan 4'ünü ise sağlıklı olarak yanlış sınıflandırmıştır. Bu çapraz hücreler yeşil olduğundan sınıflandırıcı iyi bir performans göstermiştir anlamına gelmektedir. Ayrıca gerçek sınıfın gözlemlenen örnekleri doğru bir şekilde tahmin etmiş olduğunu ifade etmektedir.



Şekil 4.11. Gözlenen örnek sayısına göre durum III için DVM sınıflandırıcısına ait karışım matrisi

Şekil 4.12’de III. durum için sağlıklı verilerinin % 100’ünü sınıflandırıcı doğru tahmin ettiğini göstermekte ve bu sınıf için Gerçek Pozitif Oranını ifade etmektedir. Nöbet anı sınıfı olan veriler 2. satırda yer almaktadır. Sınıflandırıcı bu sınıf için Gerçek Pozitif Oranını %96’ı olarak bulmaktadır. Bu 2’li sınıflandırma problemini çözmekte DVM sınıflandırıcısı III. durum için doğruluk değerini %98 olarak vermiştir.

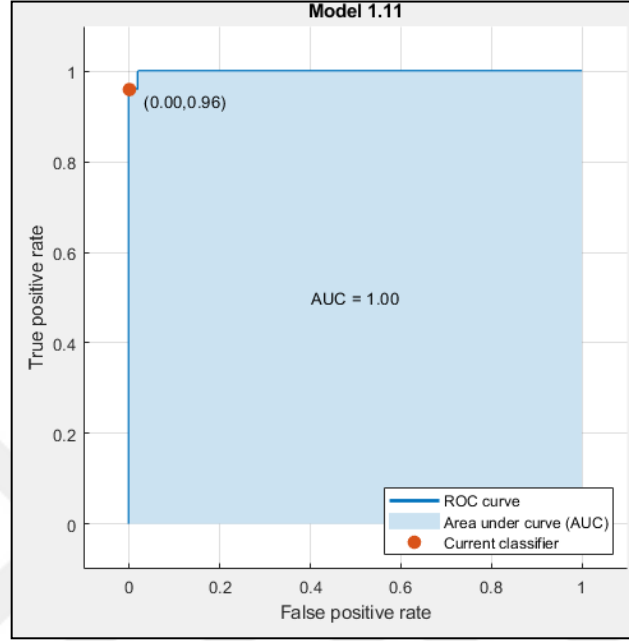


Şekil 4.12. Durum III için DVM sınıflandırıcısına ait karışım matrisi

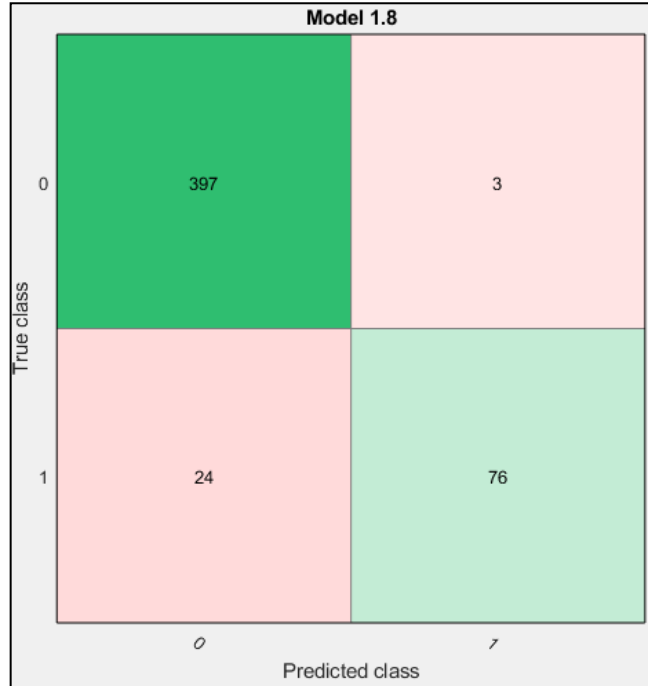
Şekil 4.13’te durum III için DVM sınıflandırıcısına ait ROC ve AUC eğrilerinde olan turuncu işaret sınıflandırıcı için hatalı pozitif oranı (FPR) ve gerçek pozitif oranı (TPR) değerleridir. Şekilde hatalı pozitif oranı 0 değeri almakta, mevcut sınıflandırıcının gözlemlenen örneklerin hiçbirini hatalı olarak pozitif sınıfa atmadığını göstermektedir. 0,96’lık gerçek pozitif oranı da sınıflandırıcının gözlemlenen örneklerin %96’sını pozitif sınıfa doğru bir şekilde atadığını ifade etmektedir. AUC 1,00 değerini almakta ve eğri altındaki alanın maksimum değeridir. Dolayısıyla bu modelin iyi bir performansa sahip sınıflandırıcı olduğunu belirtmektedir.

Şekil 4.14’te gözlenen örnek sayısına göre IV. durum için karışım matrisi yorumlanacak olursa, 1.satırda DVM sınıflandırıcısı sağlıklı sınıfına ait 400 örnekten oluşan EEG verinin 397’sini sağlıklı olarak doğru sınıflandırmıştır. Kalan 3’ünü de

nöbet anı olarak yanlış sınıflandırmıştır. Aynı şekilde model 2. satır da nöbet anı sınıfına ait 100 örnek içeren EEG kaydınının 76'sını nöbet anı olarak doğru sınıflandırmıştır, 24'dünü sağlıklı olarak yanlış sınıflandırmıştır.



Şekil 4.13. Nöbet anı pozitif sınıfı olduğunda durum III için DVM sınıflandırıcısına ait ROC ve AUC eğrileri



Şekil 4.14. Gözlenen örnek sayısına göre durum IV, DVM sınıflandırıcı, karışım matrisi

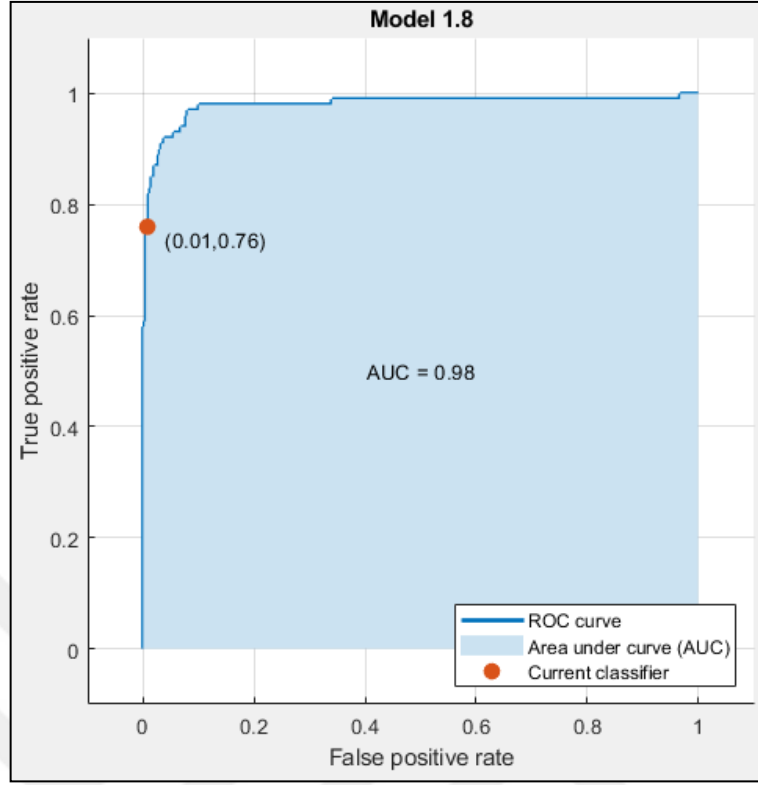
Şekil 4.15'te IV. durum için DVM modeline ait karışım matrisi verilmiştir. Bu matriste sınıflandırıcı 1. satırdaki sağlıklı EEG'lerden %1'i yanlış ve nöbet anı olarak sınıflandırmıştır. Bu sınıf için Yanlış Sınıf Oranını ifade etmektedir. Nöbet anı için EEG'lerin %24'dünü sağlıklı olarak yanlış sınıflandırılmıştır ve bu sınıf için Yanlış Sınıf Oranıdır.



Şekil 4.15. Durum IV, DVM sınıflandırıcı, karışım matrisi

Sınıflandırıcı sağlıklı EEG kayıtlarının %99'unu doğru bir şekilde sınıflandırdığını göstermekte ve bu sınıf için Gerçek Pozitif Oranını ifade etmektedir. 2. satırda nöbet anı sınıfı olan EEG kayıtlarıdır ve model %76'sını doğru bir şekilde sınıflandırdığını göstermektedir. DVM modeli IV. durum için verdiği doğruluk değeri %94,6'dır.

Durum IV için DVM sınıflandırıcısına ait ROC ve AUC eğrileri olan Şekil 4.16'nın üzerindeki işaretleyici sınıflandırıcı için hatalı pozitif oranı (FPR) ve gerçek pozitif oranı (TPR) değerlerini göstermektedir. Şekildeki 0,01'lik hatalı pozitif oranı, mevcut sınıflandırıcının gözlemlenen örneklerin %1'ini hatalı olarak pozitif sınıfa atadığını göstermektedir. 0,76'lık gerçek pozitif oranı ise sınıflandırıcının gözlemlenen örneklerin %76'sını pozitif sınıfa doğru bir şekilde atadığını göstermektedir. Şekildeki AUC değeri 0,98'dir.



Şekil 4.16. Nöbet anı pozitif sınıfı olduğunda DVM sınıflandırıcı, durum IV için ROC ve AUC eğrileri

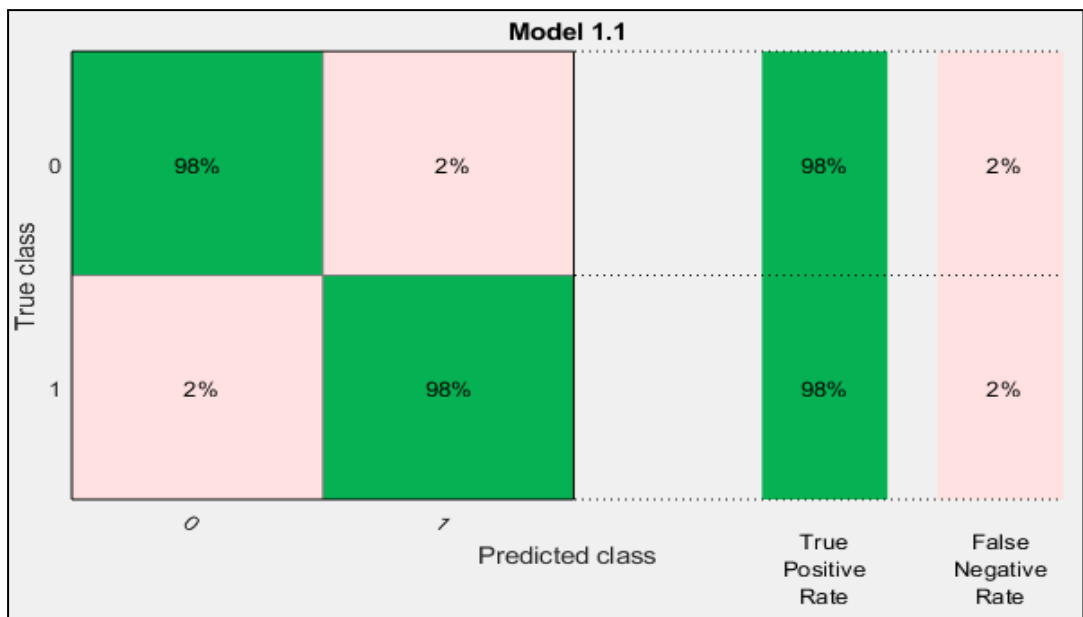
Şekil 4.17’de 1.satırda IV. durum için KA sınıflandırıcısı sağlıklı sınıfına ait 100 örnekten 98’ini sağlıklı olarak doğru sınıflandırmıştır ve kalan 2’sini nöbet anı olarak yanlış sınıflandırmıştır. Aynı şekilde sınıflandırıcı 2. satırda nöbet anı sınıfına ait 100 örneğin 98’ini nöbet anı olarak doğru sınıflandırmıştır, kalan 2’sini ise sağlıklı olarak yanlış sınıflandırmıştır. Çapraz hücreler yeşil dolayısıyla gerçek sınıfın gözlemlenen örnekleri doğru bir şekilde sınıflandırmış olduğunu belirtmektedir.

Şekil 4.8’de sağlıklı EEG kayıtlarındaki bazı örnekleri model yanlış sınıflandırmıştır. Yani %2’sini yanlış bir şekilde nöbet anı olarak sınıflandırmıştır. En son sütündeki %2 oranı, bu sınıftaki yanlış sınıflandırılmış Yanlış Sınıf Oranını ifade etmektedir. Nöbet anı EEG’lerinin %2’si sağlıklı olarak yanlış sınıflandırılmıştır. Yanlış Sınıf Oranını da %2 olarak verilmiştir.



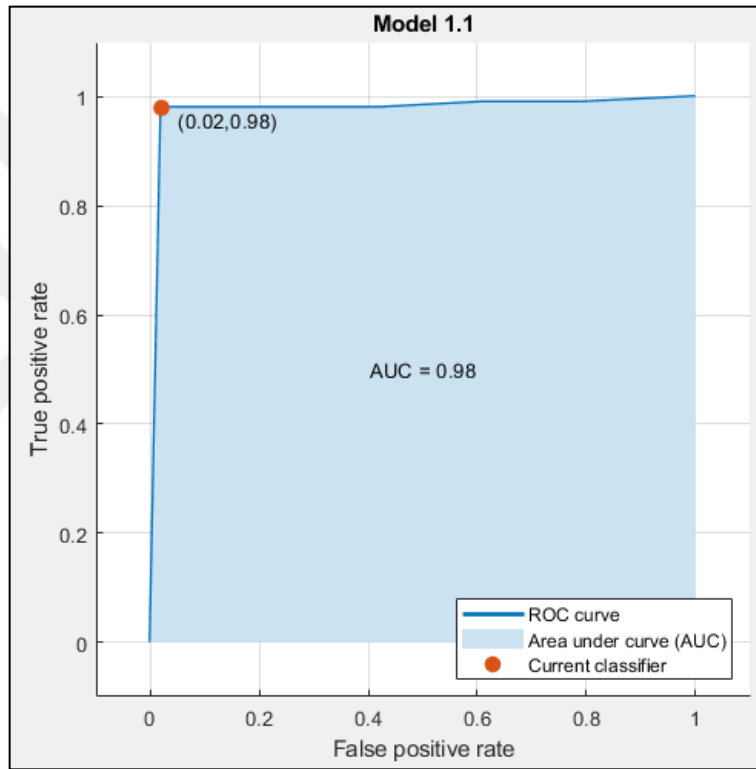
Şekil 4.17. Gözlenen örnek sayısına göre KA, durum V için karışım matrisi

Şekil 4.18’de en üstteki satırda IV.durum için sağlıklı EEG verilerinin % 98’ini sınıflandırıcı doğru tahmin ettiğini göstermekte ve bu sınıf için Gerçek Pozitif Oranını ifade etmektedir. Nöbet anı sınıfı olan EEG kayıtları da 2. satırda yer almaktadır. Sınıflandırıcı bu sınıf için Gerçek Pozitif Oranını %98 olarak göstermektedir. KA sınıflandırıcı IV. durum için doğruluk değerini %98 olarak hesaplamıştır.



Şekil 4.18. KA, durum V için karışım matrisi

Şekil 4.19'da IV. durum için KA sınıflandırıcısına ait ROC ve AUC eğrilerindeki turuncu işaret bu sınıflandırıcı için hatalı pozitif oranı ve gerçek pozitif oranıdır. Şekildeki 0,02 hatalı pozitif oranı, mevcut sınıflandırıcının gözlemlenen örneklerin %2'sini hatalı olarak pozitif sınıfa atadığını göstermektedir. 0,98'lik gerçek pozitif oranı da sınıflandırıcının gözlemlenen örneklerin %98'ini pozitif sınıfa doğru bir şekilde atadığını ifade etmektedir. AUC 0,98 değerini almakta ve eğri altındaki alan büyük olduğundan dolayı KA bu model iyi bir performansa sahip olduğunu göstermektedir.



Şekil 4.19. Nöbet anı pozitif sınıfı olduğunda V. durum için KA modeline ait ROC ve AUC eğrileri

AKA yöntemini kullanarak elde edilen Karışım matrisleri, ROC ve AUC eğrileri Şekil 4.20-4.34'te verilmiştir.

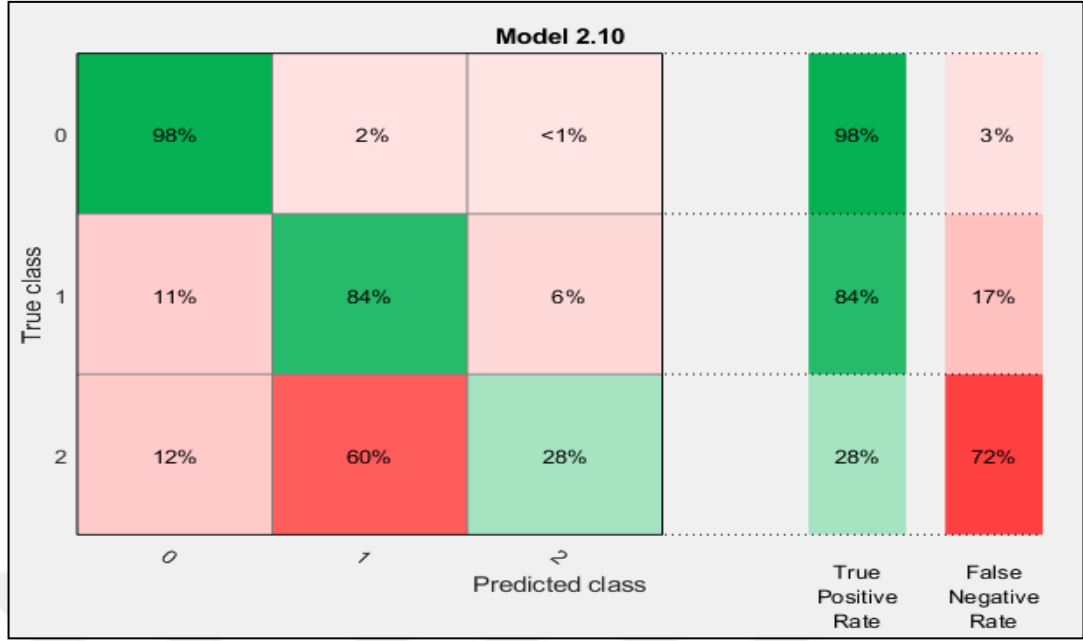
Şekil 4.20'da 1.satırda DVM sınıflandırıcısı sağlıklı sınıfına ait 200 örnekten oluşan EEG verinin 195'ini sağlıklı olarak doğru tahmin etmiştir. Kalan 4'dünü nöbet öncesi ve 1'ini de nöbet anı olarak yanlış tahmin etmiştir. Aynı şekilde 2. satır yorumlanacak olursa, nöbet öncesi sınıfına ait 200 örnek içeren EEG kaydınının 167'sini model nöbet öncesi olarak doğru tahmin etmişse, 21'ini sağlıklı ve 12'sini

nöbet anı olarak yanlış tahmin etmiştir. Üçüncü satırda 100 örnekten oluşan nöbet anına ait EEG'nin 28'ini model doğru tahmin etmiştir. Diğer kalan 12'sini sağlıklı ve 60 örneği sınıflandırıcı nöbet öncesi olarak yanlış tahmin edip, zayıf bir performans göstermiştir.

		Model 2.10		
		0	1	2
True class	0	195	4	1
	1	21	167	12
	2	12	60	28
		Predicted class		
		0	1	2

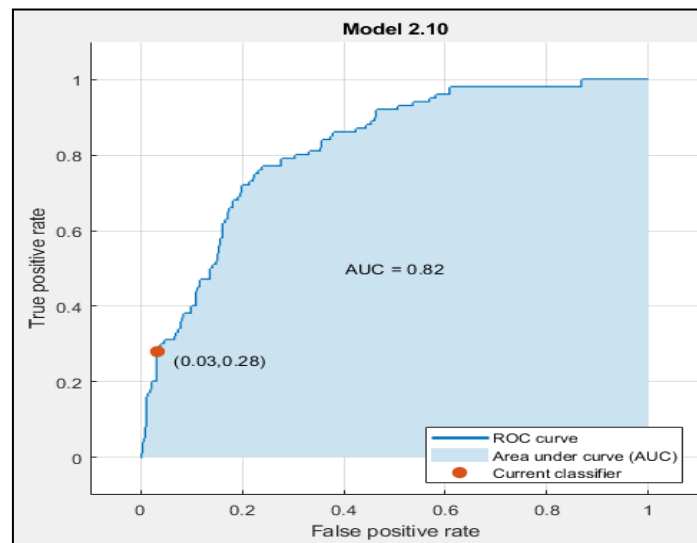
Şekil 4.20. Gözlenen örnek sayısına göre durum I, DVM modeli, karışım matrisi

Şekil 4.21'de I.durum için elde edilen DVM modeline ait karışım matrisi verilmiştir. Sağlıklı EEG kayıtları 1. satırda yer almaktadır. Sınıflandırıcı sağlıklı verinin %2'sini yanlış olarak nöbet öncesi ve %1'den az miktarını da nöbet anı olarak yanlış sınıflandırmıştır. Son sütündeki açık pembe ile gösterilen %3 oranı, bu sınıftaki yanlış sınıflandırılmış Yanlış Sınıf Oranıdır. Nöbet öncesi için verilerin %11'i sağlıklı ve %6'sı nöbet anı olarak yanlış tahmin edilmiştir. Yanlış Sınıf Oranı %17'dir. Nöbet anı verilerinin %12'si sağlıklı ve %60'ı nöbet öncesi olarak yanlış sınıflandırılmıştır. Dolayısıyla Yanlış Sınıf Oranını %72 olarak hesaplanmıştır. Hücredeki kırmızı renk bu sınıflandırıcının zayıf bir performans gösterdiğini ifade etmektedir.



Şekil 4.21. Durum I, DVM modeli, karışım matrisi

En üst sırada, sağlıklı verilerinin %98'i doğru bir şekilde sınıflandırıldığı göstermektedir. Aynı şekilde bu değer bu sınıf için Gerçek Pozitif Oranını ifade etmektedir. 2. satırda nöbet öncesi sınıfı olan verilerin %84'dü model tarafından doğru bir şekilde sınıflandırıldığını göstermektedir. 3. satırda nöbet anı sınıfı olan verilerin ancak %28 doğru bir şekilde tahmin edilmiştir. DVM sınıflandırıcı çoklu sınıflandırma problemini çözmekte bu durum için doğruluk değerini %78 olarak elde etmiştir.



Şekil 4.22. Nöbet anı pozitif sınıfı olduğunda durum I için DVM modeli, ROC ve AUC eğrileri

Durum I için DVM sınıflandırıcısına ait ROC ve AUC eğrileri olan Şekil 4.22'nin üzerindeki işaretleyici hatalı pozitif oranını ve gerçek pozitif oranını ifade etmektedir. Şekildeki 0,03'lük hatalı pozitif oranı, mevcut sınıflandırıcının gözlemlenen örneklerin %3'ünü hatalı olarak pozitif sınıfa atadığını gösterir. 0,28'lik gerçek pozitif oranı ise sınıflandırıcının gözlemlenen örneklerin %28'ini pozitif sınıfa doğru şekilde atadığını göstermektedir. AUC 0,82 değerini almakta bu seçilen modelin performansını daha önceki durumlara göre zayıf bir sonuç olduğunu ifade etmektedir.

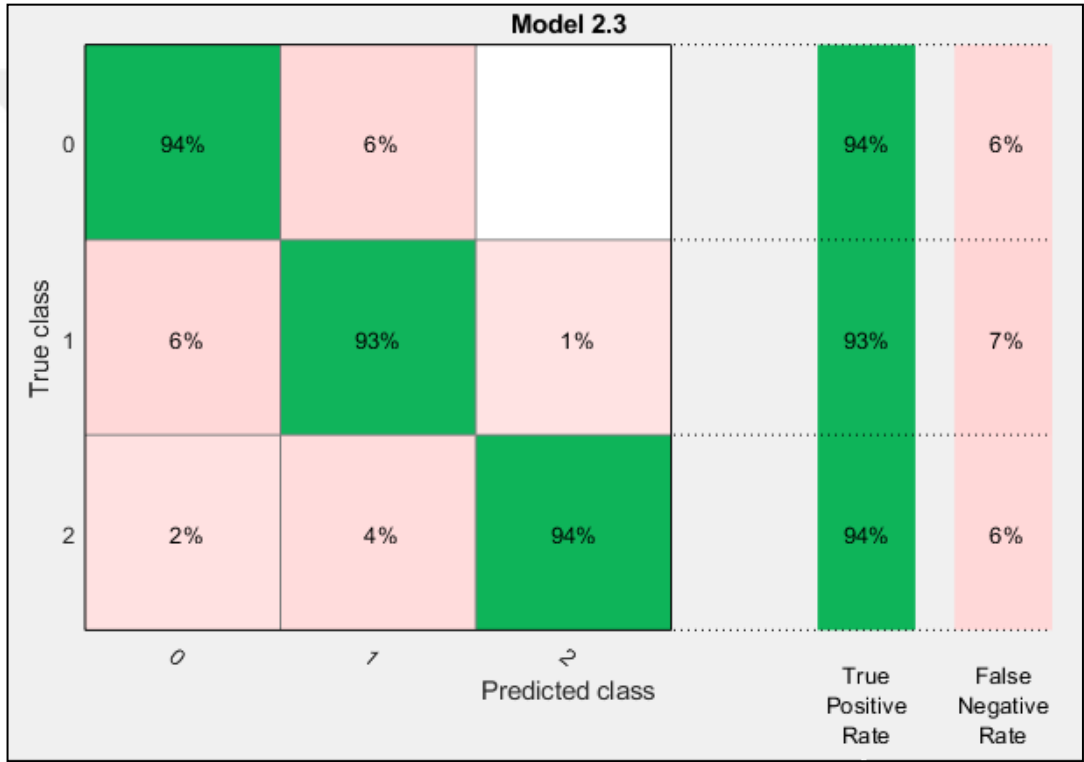
		Model 2.3		
		0	1	2
True class	0	94	6	
	1	6	93	1
	2	2	4	94
		0	1	2
		Predicted class		

Şekil 4.23. Gözlenen örnek sayısına göre durum II, KA modeli, karışım matrisi

Şekil 4.23'te II. durum için KA sınıflandırıcısı 1.satırda sağlıklı sınıfına ait 100 örnekten 94 örneği sağlıklı olarak doğru tahmin etmiş ve kalan 6 örneği nöbet öncesi olarak yanlış tahmin etmiştir. Aynı şekilde 2. satır da verilen nöbet öncesi sınıfına ait 100 örneğin 93 örneği nöbet öncesi olarak doğru tahmin edilmiştir. Kalan 6 örneği sağlıklı ve 1 örneği de nöbet anı olarak yanlış sınıflandırmıştır. Son satırdaki 100 örnekten oluşan nöbet anı sınıfınının 94 örneğini sınıflandırıcı doğru tahmin etmiştir. Kalan 2 örneği sağlıklı ve 4 örneği de nöbet öncesi olarak yanlış sınıflandırmıştır. Çapraz hücreler yeşil olduğundan sınıflandırıcı iyi bir performans göstermiş

anlamına gelmektedir. Dolayısıyla gerçek sınıfın gözlemlenen örnekleri doğru bir şekilde sınıflandırmış olduğunu da ifade etmektedir.

Şekil 4.24'te II.durum için elde edilen KA modeline ait karışım matrisidir. Sağlıklı verilerin %6'sını nöbet öncesi olarak yanlış sınıflandırmıştır. Bu değer de sınıftaki Yanlış Sınıf Oranını ifade etmektedir. Nöbet öncesi için verilerin %6'sını sağlıklı ve %1'ini nöbet anı olarak yanlış sınıflandırmıştır. Dolayısıyla Yanlış Sınıf Oranı %7'dir. Model nöbet anı verilerinin %2'sini sağlıklı ve %4'dünü nöbet öncesi olarak yanlış sınıflandırmıştır. Demek ki bu sınıf için Yanlış Sınıf Oranı %6'dır.

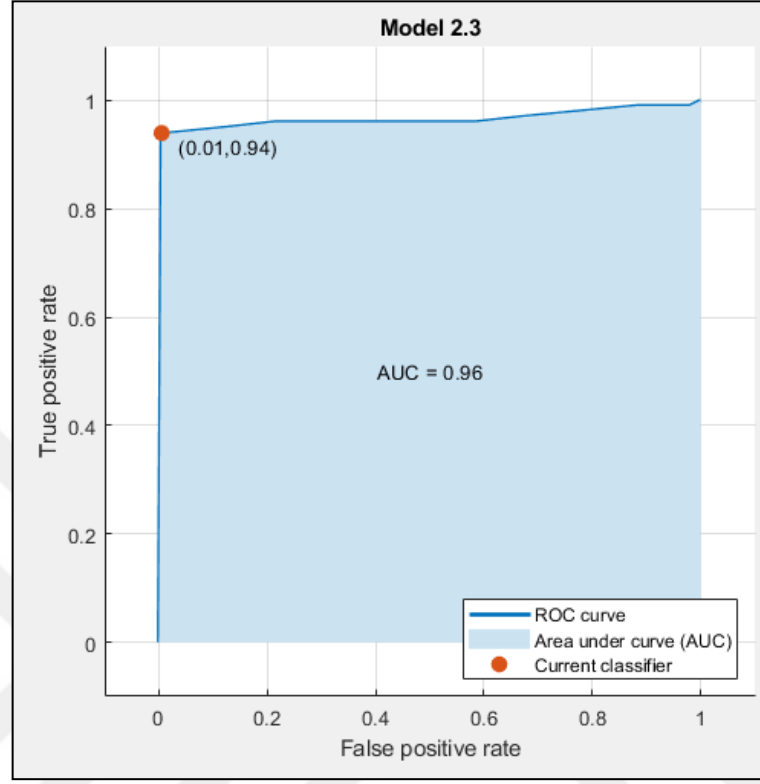


Şekil 4.24. Durum II, KA modeli, karışım matrisi

Yeşil renkli çapraz hücrelerden ortaya çıkarak, sağlıklı verilerin %94'i doğru bir şekilde sınıflandırılmıştır. Aynı şekilde 2. satırda nöbet öncesi sınıfı olan verilerin %93'ü doğru bir şekilde sınıflandırılmıştır. Nöbet anı sınıfı olan son satırdaki verilerin %94'dünü sınıflandırıcı doğru tahmin etmiştir. Bu değerler her bir sınıf için Gerçek Pozitif Oranlarını ifade etmektedir. KA sınıflandırıcısı çoklu sınıflandırma problemi için doğruluk değerini %93,7 olarak hesaplamıştır.

Şekil 4.25'te ROC eğrisindeki %1 değeri sınıflandırıcının gözlemlenen örneklerin hatalı olarak pozitif sınıfa atadığını, %94'dünü ise pozitif sınıfa doğru şekilde

atadığını göstermektedir. AUC 0,96 demek ki modeli iyi bir performans göstermektedir.

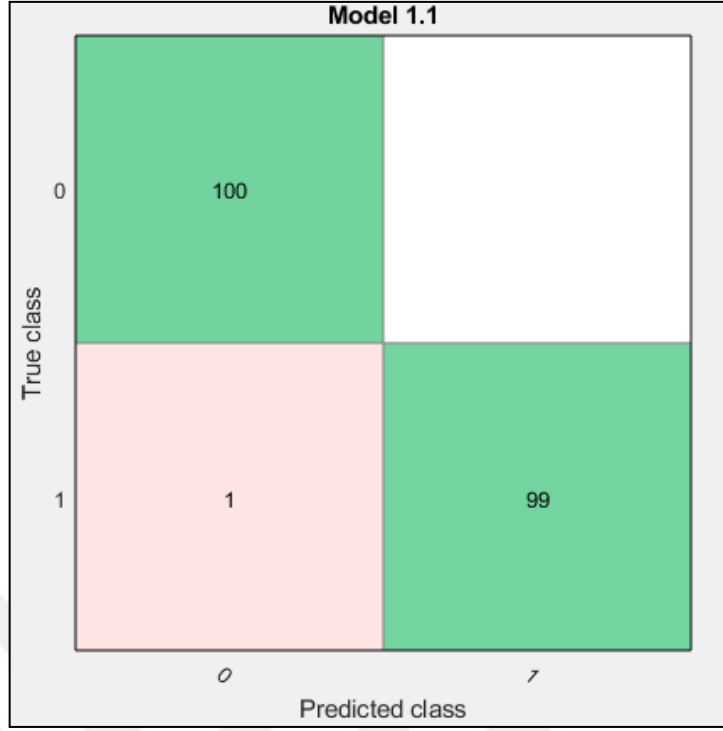


Şekil 4.25. Nöbet anı pozitif sınıfı olduğunda durum II için KA modeli, ROC ve AUC eğrileri

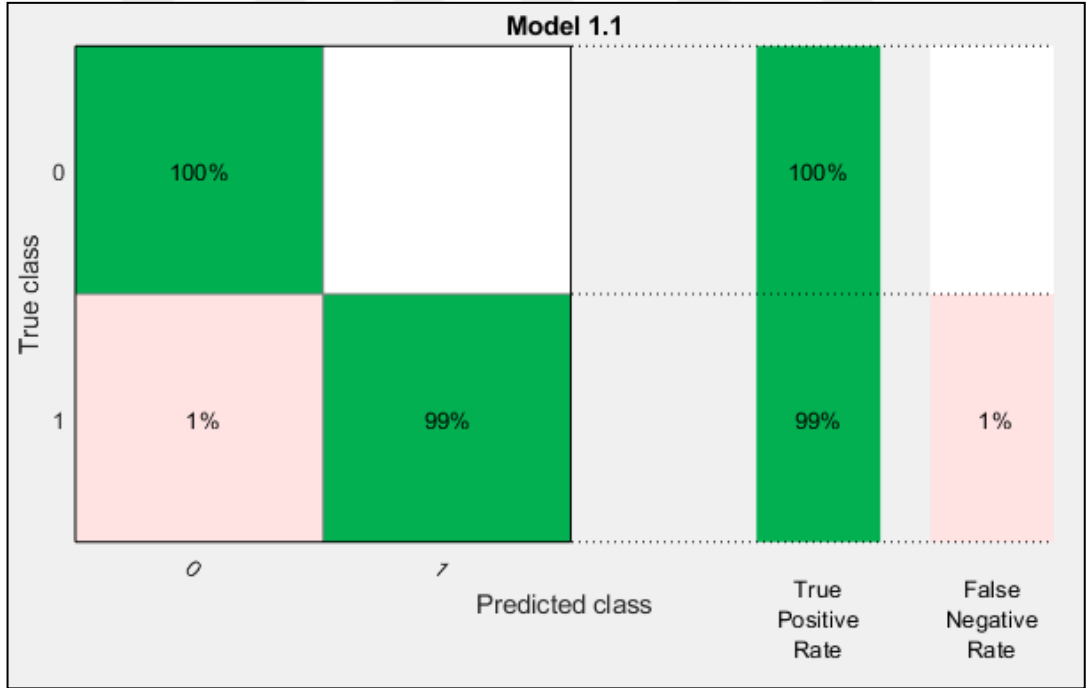
Şekil 4.26 ve Şekil 4.27'de III.durum için KA modeline ait karışım matrisi verilmiştir. Sağlıklı sınıf için 100 örneğin %100'ü doğrudur. Nöbet anı sınıfı için 100 örneğin %99'u doğru tahmin edilmiştir. Doğruluk değeri de %99,5 olarak daha önceki ADD yönteminden iyi bir sonuç göstermektedir.

Şekil 4.28'de ROC eğrisinde sınıflandırıcının gözlemlenen örneklerin %99'unu pozitif sınıfa doğru şekilde atadığını göstermektedir. AUC 0,99 eğri altındaki alan büyük olduğundan dolayı seçilen modelin iyi bir performansa sahip sınıflandırıcı olduğunu belirtmektedir.

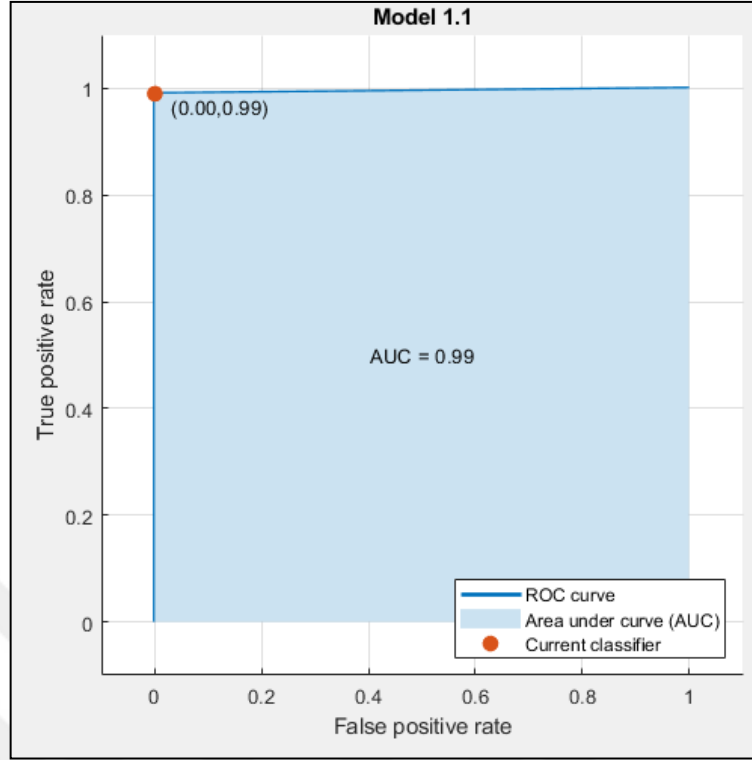
Şekil 4.29 ve Şekil 4.30'da IV.durum için KA modeline ait karışım matrisi verilmiştir. Sağlıklı sınıf için 398 örnek ve neredeyse %100 doğrudur. Nöbet anı sınıfı için 96 örneğin %96'sını doğru bir şekilde tahmin edilmiştir. Doğruluk değeri de %98,8 olarak elde edilmiştir. ADD yönteminden daha iyi bir sonuç göstermektedir.



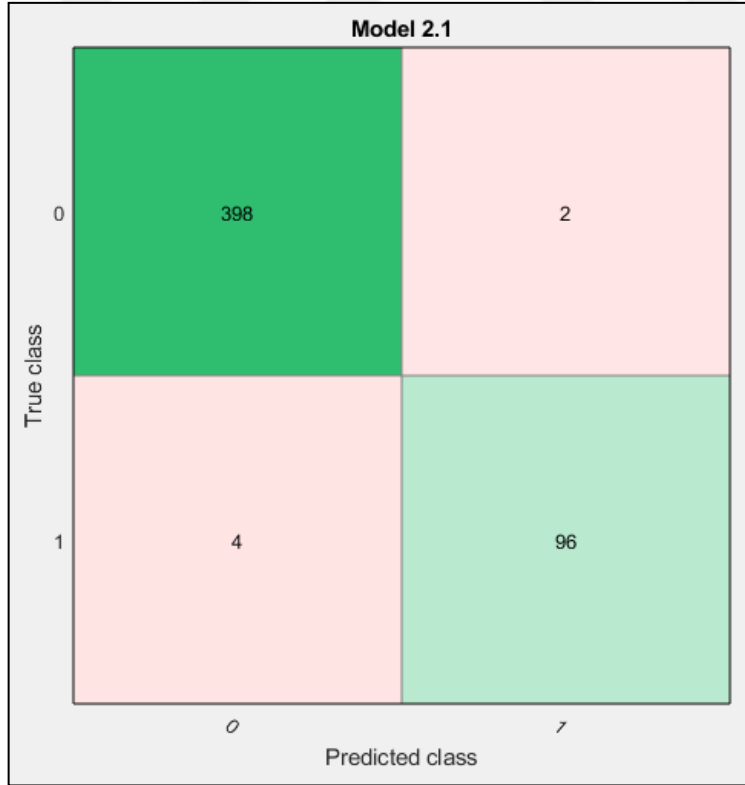
Şekil 4.26. Gözlenen örnek sayısına göre III. durum için KA modeli, karışım matrisi



Şekil 4.27. Durum III, KA modeli, karışım matrisi



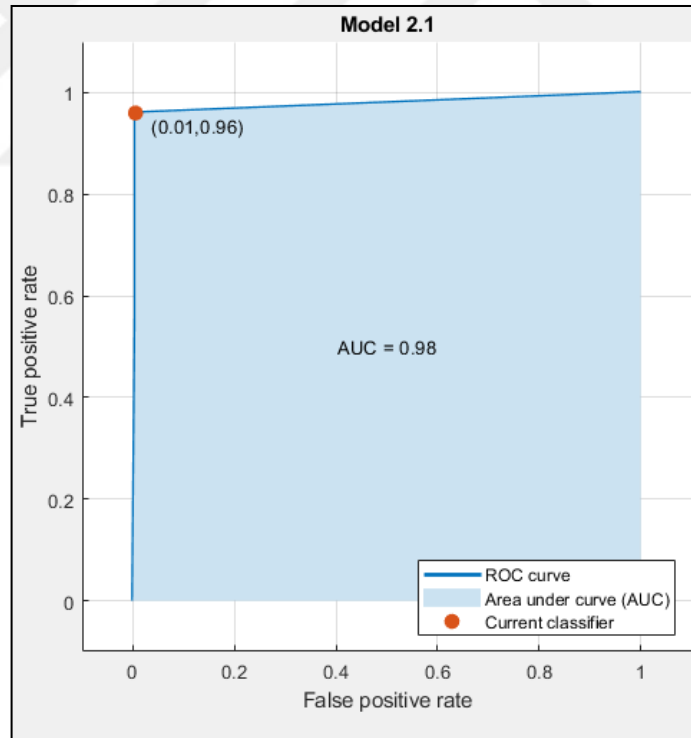
Şekil 4.28. Nöbet anı pozitif sınıfı olduğunda durum III için KA modeli, ROC ve AUC eğrileri



Şekil 4.29. Gözlenen örnek sayısına göre durum IV, KA modeli, karışım matrisi



Şekil 4.30. Durum IV, KA modeli, karışım matrisi



Şekil 4.31. Nöbet anı pozitif sınıfı olduğunda durum IV, KA modeli, ROC ve AUC eğrileri

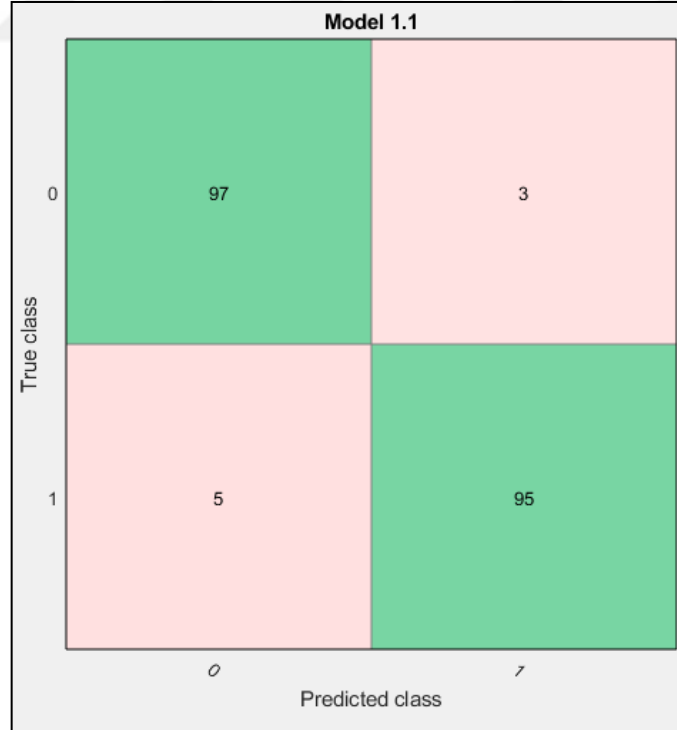
Şekil 4.32'de sınıflandırıcı 1.satırda 100 örnekten oluşan EEG verisinin 97'sini sağlıklı olarak doğru sınıflandırmıştır. Kalan 3'ünü ise nöbet anı olarak yanlış sınıflandırmıştır. Son satırda 100 örnekten oluşan nöbet anına ait EEG'nin 95'ini

model doğru tahmin etmiştir. Sınıflandırıcı kalan 5'ini sağlıklı olarak yanlış tahmin etmiştir.

Şekil 4.33'te V.durum için KA na ait karışım matrisi verilmiştir. Sağlıklı sınıf için 100 örneğin %97'si doğrudur. Nöbet anı sınıfı için 95 örneğin %95'i doğru bir şekilde tahmin edilmiştir. Doğruluk değeri de %96'dır.

Model sağlıklı EEG kayıtlarının %3'ini nöbet anı olarak yanlış sınıflandırmıştır. En son sütündeki %3 oranı, bu sınıftaki yanlış sınıflandırılmış Yanlış Sınıf Oranını ifade etmektedir. Nöbet anı için EEG'lerin %5'i sağlıklı olarak yanlış sınıflandırılmıştır. Yanlış Sınıf Oranını da %5 olarak verilmiştir.

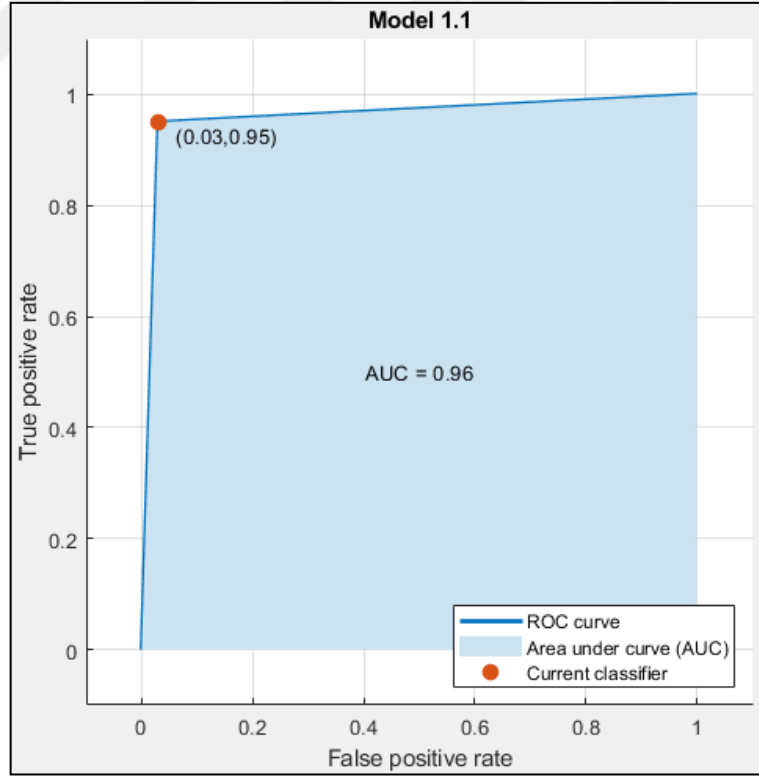
Şekil 4.34'te ROC eğrisinde 0,03'lük bir hatalı pozitif oran, mevcut sınıflandırıcının gözlemlenen örneklerin %3'ünü hatalı olarak pozitif sınıfa atadığını gösterir. 0,95'lik gerçek pozitif oran, sınıflandırıcının gözlemlenen örneklerin %95'ini pozitif sınıfa doğru şekilde atadığını ifade etmektedir. AUC değeri 0,96, demek ki sınıflandırıcının iyi bir performans gösterdiğini ifade etmektedir.



Şekil 4.32. Gözlenen örnek sayısına göre durum V, KA modeli, karışım matrisi



Şekil 4.33. Durum V, KA modeli, karışım matrisi



Şekil 4.34. Nöbet anı pozitif sınıfı olduğunda durum V, KA modeli, ROC ve AUC eğrileri

5. SONUÇLAR VE ÖNERİLER

Bir nörolog tarafından görsel olarak nöbet tespiti yapmak çeşitli nedenlerden dolayı zor olabilmektedir. Çünkü nöbet tahmininin tespiti konusunda çok az onaylanmış bilgi vardır. Ayrıca epileptik nöbetler sırasında hasta fiziksel ve zihinsel durumunun farkında olmadığı için fiziksel yaralanmalar söz konusu olabilir. Bu çalışmada epileptik nöbet teşhisinde otomatik metotların geliştirilmesi, nöbetlerin gerçekleşmeden önce engelleyecek veya hafifletecek sistemlerin tasarımına yol açmacını sağlama amaçlıdır. Bu amaçla epilepsi tespiti yapılması konusunda makine öğrenmesinden faydalanılmıştır. Makine öğrenmesinde kullanılan üç ayrı sınıflandırma yönteminin performansları incelenmiştir. Öncelikle çalışmada AKA ve ADD yöntemleri kullanılarak EEG sinyallerinden iki frekans uzayı dönüşümüne dayanarak istatistiksel moment öznitelikleri elde edilmiştir. Çıkarılan bu özniteliklere DVM, K-YK, KA yöntemleri uygulanıp sınıflandırma işlemi gerçekleştirilmiştir. Deneysel sonuçlara göre epilepsi tespiti ve tahmininde en iyi sonucu AKA yöntemi vermiştir. Sınıflandırma yöntemlerinden DVM ve KA ise yaklaşık sonuçlar ürettiği kanısına varılmıştır.

Epilepsi teşhisinin gerekli olduğu durumlarda sınıflandırma yöntemlerinin oldukça iyi performans sergilediği görülmektedir. Uzmanların olmadığı durumda EEG kayıtları kullanılarak epilepsi hastalığının teşhisi konusunda makine öğrenmesinden rahatlıkla faydalanılabilir. Bu yüzden bu kayıtların depolanması ve korunması oldukça önemlidir. Bu konuda yetkililerin gereken hassasiyeti göstermeleri önerilmektedir.

KAYNAKLAR

- [1] Yazgan E., Korürek M., *Tıp Elektronik*, İTÜ Elektrik – Elektronik Fakültesi Ofset Baskı Atölyesi, İstanbul, 323, 1996.
- [2] Smith J. R., Automatic Analysis and Detection of EEG Spikes, *IEEE Transactions on Biomedical Engineering*, 1974, **21**(1), 1-7.
- [3] Jones R. D., Dingle A. A., Fright W. R., Carroll G. J., Donaldson, I. M. Automated Detection of Epileptiform Transients in the EEG Using a Multi-Channel Expert System Approach, *Images of the Twenty-First Century. Proceedings of the Annual International Engineering in Medicine and Biology Society*, DOI: 10.1109/IEMBS.1989.95939.
- [4] Park H. S., Lee Y. H., Lee D. S., Kim S. I., Detection of Epileptiform Activity Using Wavelet and Neural Network, *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, DOI: 10.1109/IEMBS.1997.756576.
- [5] Srinivasan V., Eswaran C., Sriraam N., Artificial Neural Network Based Epileptic Detection Using Time-Domain and Frequency-Domain Features, *Journal of Medical Systems*, DOI: 10.1007/s10916-005-6133-1.
- [6] Panda R., Khobragade P. S., Jambhule P. D., Jengthe S. N., Pal P. R., Gandhi T. K., Classification of EEG Signal Using Wavelet Transform and SVM for Epileptic Seizure Diction, *International Conference on Systems in Medicine and Biology*, DOI: 10.1109/ICSMB.2010.5735413.
- [7] Murugavel A. S. M., Ramakrishnan S., Maheswari U., Sabetha B. S., Combined Seizure Index with Adaptive Multi-Class SVM for Epileptic EEG Classification, *2013 International Conference on Emerging Trends in VLSI*, DOI: 10.1109/ICEVENT.2013.6496565.
- [8] Alam S. M. S., Bhuiyan M. I. H., Detection of Seizure and Epilepsy Using Higher Order Statistics in the EMD Domain, *IEEE Journal of Biomedical and Health Informatics*, DOI: 10.1109/JBHI.2012.2237409.
- [9] Mohite N., Shastri R., Deosarkar S., Das A., Epileptic Electroencephalogram Classification, *2014 International Conference on Communication and Signal Processing*, DOI: 10.1109/ICCSP.2014.6949885.
- [10] Hamad A., Houssein E. H., Hassanien A. E., Fahmy A. A., Feature Extraction of Epilepsy EEG using Discrete Wavelet Transform, *2016 12th International Computer Engineering Conference, (ICENCO)*, DOI: 10.1109/ICENCO.2016.7856467.

- [11] Tessy E., Shanir M. P. P., Manafuddin S., Time Domain Analysis of Epileptic EEG for Seizure Detection, *2016 International Conference on Next Generation Intelligent Systems (ICNGIS)*, DOI: 10.1109/ICNGIS.2016.7854034.
- [12] Das K., Mudoir R., Analysis of EEG Signals using Empirical Mode Decomposition and Support Vector Machine, *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, DOI: 10.1109/ICPCSI.2017.8392315.
- [13] Yol S., Özdemir M. A., Akan A., Chaparro L. F., Detection of Epileptic Seizures by the Analysis of EEG Signals Using Empirical Mode Decomposition, *2018 Medical Technologies National Congress (TIPTEKNO)*, DOI: 10.1109/TIPTEKNO.2018.8596780.
- [14] Daoud H. G., Abdelhameed M. A., Bayoumi M., Automatic Epileptic Seizure Detection based on Empirical Mode Decomposition and Deep Neural Network, *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)*, DIO: 10.1109/CSPA.2018.8368709.
- [15] Li S., Zhou W., Yuan Q., Geng S., Cai D., Feature Extraction and Recognition of Ictal EEG Using EMD and SVM, *Computer in Biology and Medicine Journal*, 2013, **43**(7), 807-816.
- [16] Gümüş İ., EKG Sinyallerinin Wavelet Analizi, Lisans Tezi, Uludağ Üniversitesi, Mühendislik Mimarlık Fakültesi Elektronik Mühendisliği Bölümü, Bursa, 2003.
- [17] Toprak B. İ., EEG Sinyallerinin Dalgacık Dönüşümü ve Yapay Sinir Ağları ile Analizi, Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Isparta, 2007, 200549.
- [18] Kim D., Introduction to EMD (Empirical Mode Decomposition) with Application to a Scientific Data, <http://stat.snu.ac.kr/heeseok/html/Talk/EMDintroSeminar.pdf> (Ziyaret tarihi: 5 Haziran 2019).
- [19] Çomak E., Destek Vektör Makineleri Çoklu Sınıf Problemleri İçin Çözüm Önerileri, Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya, 2004, 153931.
- [20] Burges C. J. C., Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining Knowledge Discovery*, 1998, **2**(2), 121-167.
- [21] Alpaydin E., *Introduction to Machine Learning*, 3rd ed., MIT Press, Massachusetts, 2004.
- [22] Demirci D. A., Destek Vektör Makineleri ile Karakter Tanıma, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2007, 201495.

- [23] Guo L., Rivero D., Dorado J., Munteanu C. R., Pazos A., Automatic Feature Extraction Using Genetic Programming: an Application to Epileptic EEG Classification, *Expert Systems with Applications*, 2011, **38**(8), 10425-10436.
- [24] URL-1: <http://www.bem.fi/book/13/13.htm> (Ziyaret tarihi: 13 Mayıs 2019).
- [25] URL-2: https://forecasting.svetunkov.ru/forecasting_toolbox/data-analysis-stat/, (Ziyaret tarihi: 14 Mayıs 2019).
- [26] URL-3: <https://medium.com/@k.ulgen90/makine-ogrenimi-bolum-2-6d6d12018e1> (Ziyaret tarihi: 5 Haziran 2019).
- [27] URL-4: <http://erdoganb.com/2017/07/karar-agaclari-decision-trees-ile-veri-si-niflandirma/> (Ziyaret tarihi: 5 Haziran 2019).
- [28] URL-5: <https://medium.com/@raghaviadoni/evaluation-metrics-i-precision-recall-and-f1-score-3ec25e9fb5d3> (Ziyaret tarihi: 16 Mayıs 2019).
- [29] URL-6: <https://www.mathworks.com/help/stats/assess-classifier-performance.html#bun1ao7> (Ziyaret tarihi: 20 Haziran 2019).

KİŞİSEL YAYIN VE ESERLER

- [1] **Zhunis A.**, Çelebi T. A., Destek Vektör Makineleri Kullanılarak Epileptik Elektroensefalogram Tespiti, 2. *Uluslararası Öğrenciler Fen Bilimleri Kongresi*, İzmir, Türkiye, 4-5 Mayıs 2018.
- [2] **Zhunis A.**, Çelebi T. A., EEG İşaretlerinde Epilepsi Nöbet Tahmini ve Tespiti için Ampirik Kip Ayrışımı ve Ayrık Dalgacık Dönüşümü Yöntemlerinin Performans Karşılaştırılması, *Uluslararası Marmara Fen Bilimleri Kongresi*, Kocaeli, Türkiye, 26-28 Nisan 2019.



ÖZGEÇMİŞ

1992 yılında Sarkand'ta (Kazakistan) doğdu. İlk ve ortaöğrenimini Almatı 139 nolu okulda lise eğitimini ise 175 nolu lisede tamamladı. Uluslararası Bilgi Teknolojileri Üniversitesinde (IITU) Radyo, Elektronik ve Haberleşme Mühendislik Fakültesi Bölümü'nü 2014 yılında bitirdi. 2015 yılında Kocaeli Üniversitesi Fen Bilimleri Enstitüsü'nde Yüksek Lisans programına başladı.

