

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**MELEZ SINIFLANDIRMA YAKLAŞIMI İLE EEG İŞARETLERİNDEN
EPİLEPTİK DÖNEMLERİN ALGILANMASI**

YÜKSEK LİSANS TEZİ

Bilgisayar Müh. Funda KUTLU

**Ocak 2014
TRABZON**

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

MELEZ SINIFLANDIRMA YAKLAŞIMI İLE EEG İŞARETLERİNDEN
EPİLEPTİK DÖNEMLERİN ALGILANMASI

Bilgisayar Mühendisi Funda KUTLU

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

Tezin Enstitüye Verildiği Tarih : 13.12.2013
Tezin Savunma Tarihi : 03.01.2014

Tez Danışmanı : Doç. Dr. Cemal KÖSE

Trabzon 2014

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalında

Funda Kutlu tarafından hazırlanan

**MELEZ SINIFLANDIRMA YAKLAŞIMI İLE EEG İŞARETLERİNDEN
EPİLEPTİK DÖNEMLERİN ALGILANMASI**

**başlıklı bu çalışma, Enstitü Yönetim Kurulunun 17 / 12 / 2013 gün ve 1534
sayılı kararıyla oluşturulan jüri tarafından yapılan sınavda**

YÜKSEK LİSANS TEZİ

olarak kabul edilmiştir.

Jüri Üyeleri

Başkan : Doç. Dr. Cemal KÖSE



Jüri : Yrd. Doç. Dr. Bekir DİZDAROĞLU



Jüri : Yrd. Doç. Dr. Önder AYDEMİR



Prof. Dr. Sadettin KORKMAZ

Enstitü Müdürü

ÖNSÖZ

Bilgisayarlı sistemlerle hastalıkların ön tanısının gerçekleştirilmesi, hem zamandan tasarruf sağlamak hem de hastalıkların yanlış tanısını azaltmak açısından son dönemlerde popüler olan konulardan biri olmuştur. Bu amaçla çalışmada, epilepsi hastalığının ön tanısının gerçekleşmesi için sağlıklı, nöbet öncesi ve epileptik nöbet dönemlerinin ayrıştırılmasını sağlayan bir otomatik tanı sistemi geliştirilmiştir. Önerilen bu sistem, verilerin zenginleştirilmesi ve gerçek zamanlı hale getirilmesi durumunda, biyomedikal alanına katkı sağlayan bir sistem olacaktır.

Yüksek Lisans eğitimim ve tez çalışmasını yürütme süresince engin tecrübe ve bilgileriyle yol gösteren, zaman ayıran Sayın Tez Danışmanım Doç. Dr. Cemal KÖSE'ye teşekkürlerimi sunarım. Edindikleri tecrübelerini ve fikirlerini eksik etmeyen hocalarım ve meslektaşlarım Sn. Arş. Gör. Selda BAYRAK, Sn. Dr. Murat AYKUT ve çalışmamda katkısı olan mesai arkadaşlarıma; Sn. Yrd. Doç. Dr. Önder AYDEMİR ve Sn. Dr. Mehmet ÖZTÜRK'e; çalışmam süresince sabır, destek ve yardımlarını esirgemeyen nişanlım Kontrol Müh. Cemal ONAY'a, aileme ve tüm dostlarıma teşekkürü borç bilirim.

Funda KUTLU
Trabzon 2014

TEZ BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “Melez Sınıflandırma Yaklaşımı ile EEG İşaretlerinden Epileptik Dönemlerin Algılanması” başlıklı bu çalışmayı baştan sona kadar danışmanım Doç. Dr. Cemal KÖSE’ nin sorumluluğunda tamamladığımı, verileri kendim topladığımı, deneyleri ve analizleri ilgili laboratuvarlarda yaptığımı, başka kaynaklardan aldığım bilgileri metin içerisinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma süresince bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 03/01/2014



Funda KUTLU

İÇİNDEKİLER

ÖNSÖZ.....	III
TEZ BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET	VII
SUMMARY	VIII
ŞEKİLLER DİZİNİ	IX
TABLolar DİZİNİ.....	XI
SEMBOLLER DİZİNİ	XII
1. GENEL BİLGİLER.....	1
1.1. Giriş.....	1
1.2. Beyin Hücresi ve Çalışması.....	3
1.3. Otomatik Tanı Sistemleri.....	4
1.4. EEG İşaretlerinin Özellikleri ve Ölçülmesi.....	5
1.5. Epilepsi Hastalığının Tanımı ve Özellikleri	7
1.6. Tezin Genel Yapısı ve Amacı.....	8
1.7. Öznitelik Çıkarımı	9
1.7.1. Zaman Uzayında Öznitelik Çıkarımı.....	9
1.7.1.1. Hjorth Betimleyicileri.....	9
1.7.1.2. Tekrarlılık Grafiği (TG) ve Tekrarlılık Ölçme Analizi (TÖA)	10
1.7.1.2.1. Tekrarlılık Grafiği (TG).....	10
1.7.1.2.2. Tekrarlılık Ölçme Analizi (TÖA).....	11
1.7.1.3. Sıfır Geçiş Oranı.....	13
1.7.1.4. Hurst Üstel Katsayısı.....	14
1.7.1.5. Entropi.....	15
1.8. Sınıflandırma Yöntemleri	16
1.8.1. Yapay Sinir Ağları.....	16
1.8.1.1. Yapay Sinir Hücresi (Nöron).....	18
1.8.1.2. Algılayıcı	19
1.8.1.3. Çok Katmanlı Yapay Sinir Ağları	22

1.8.1.4.	ÇKYSA'nın Öğrenme Kuralı	22
1.8.1.4.1.	İleri Doğru Hesaplama.....	23
1.8.1.4.2.	Geriye Doğru Hesaplama	24
1.8.1.5.	ÇKYSA'nın Çalışma Prensipleri	26
1.8.2.	k-En Yakın Komşuluk.....	27
1.8.3.	Naive Bayes Sınıflandırıcı.....	28
1.8.4.	Destek Vektör Makineleri	32
1.8.4.1.	Doğrusal Ayrılabilen Veriler İçin DVM	33
1.8.4.2.	Doğrusal Ayrılamayan Veriler İçin DVM.....	36
1.8.4.3.	DVM ile Çoklu Sınıflandırma	38
2.	YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME	40
2.1.	Çalışmanın Yapısı ve Adımları	40
2.2.	Kullanılan Veri Seti	40
2.3.	Özniteliklerin Elde Edilmesi	42
2.3.1.	Hjorth Betimleyicilerinin Uygulanması	42
2.3.2.	Tekrarlılık Ölçme Analizinin Uygulanması	44
2.3.3.	Sıfır Geçiş Oranının Elde Edilmesi.	45
2.3.4.	Hurst Üstel Katsayısının Elde Edilmesi	46
2.3.5.	Entropi Değerinin Elde Edilmesi.....	47
2.4.	Çapraz Doğrulama.....	48
2.5.	Sınıflandırma Sonuçlarının Elde Edilmesi ve Karşılaştırılması.	50
2.5.1.	Sınıflandırıcı Performanslarının Elde Edilmesi.....	53
2.5.2.	Melez Yaklaşımlar ve Borda Sayısı ile Sınıflandırıcıların Ağırlıklandırılması	55
2.5.3.	Tekil ve Melez Sınıflandırma Sonuçları.....	57
2.5.4.	Sonuçların Karşılaştırılması	61
3.	SONUÇLAR VE ÖNERİLER.....	65
4.	KAYNAKLAR.....	67
ÖZGEÇMİŞ		

Yüksek Lisans Tezi

ÖZET

MELEZ SINIFLANDIRMA YAKLAŞIMI İLE EEG İŞARETLERİNDEN EPİLEPTİK DÖNEMLERİN ALGILANMASI

Funda KUTLU
Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Doç. Dr. Cemal KÖSE
2014, 73 Sayfa

Günümüzde otomatik tanı sistemlerinin tıbbi alanlarda gittikçe yaygınlaşmasıyla, hastalıkların tanısı için harcanan süre ve yanlış tanımlar azalmıştır. Bu sistemler aynı zamanda hastalığın tedavisinde uzmanlara da yol gösterir nitelikte olmuştur. Daha önceki teşhislere yönelik verilerin saklanması ve yeniden yorumlanması adına kolaylık sağladığından, bu sistemlerin geliştirilmesinin ve desteklenmesinin önemi artmıştır.

Tez çalışması, epileptik nöbetin teşhisi amaçlı sağlıklı, nöbet öncesi ve nöbet dönemi verilerinin sınıflandırılması ve bu sınıflandırma başarılarının karşılaştırmalı olarak ele alınmasını içermektedir. Bunun için daha önce hastalık teşhisi ya da farklı amaçlarla EEG analizinde kullanılmış olan Hjorth betimleyicileri, tekrarlılık ölçme analizi, Shannon entropi, Hurst üstel katsayısı ve sıfır geçiş oranı yöntemlerinden faydalanılmıştır. Elde edilen öznitelikler, k-en yakın komşuluk, Naive Bayes, yapay sinir ağları ve destek vektör makineleri ile önce tekli daha sonra performansı artırmak için melez sınıflandırma yapılmıştır. Çalışmada, hem epileptik nöbetin tespiti, hem de hastalıklı ve sağlıklı bireylerin ayırt edilmesine ilişkin veri gruplamaları yapılmış ve 7 adet alt veri kümesi oluşturulmuştur. Önerilen sistemin performansı doğruluk, duyarlılık ve özgünlük ölçütleri ile, oluşturulan her bir veri kümesi için belirlenmiştir. Bu durumda %97.5 ila %100 arasında sınıflandırma doğruluğu bulunmuş ve sonuçlar literatürdeki çalışmalarla kıyaslandığında tez çalışmasında önemli ilerlemelerin elde edildiği görülmüştür.

Anahtar Kelimeler: Epileptik Nöbet, Hjorth, Entropi, Tekrarlılık Ölçme Grafiği, Hurst Üstel Katsayısı, Melez Sınıflandırıcılar

Master Thesis

SUMMARY

DETECTION OF EPILEPTIC SEIZURES FROM THE EEG SIGNALS
BY USING A HYBRID CLASSIFICATION APPROACH

Funda KUTLU
Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Computer Engineering Graduate Program
Supervisor: Assoc. Prof. Dr. Cemal KÖSE
2014, 73 Pages

Nowadays, the implementation of computerized systems in the medical field is quite popular. For this reason, the time taken for diagnosis of diseases and incorrect diagnoses have been decreased. These systems are also employed by the experts to improve treatment of illnesses. Because of these systems make storing and reinterpretation of diagnostic data more easier, the importance of development and support of these systems has increased.

In this thesis, the data of healthy, pre-seizure and seizure period have been classified to diagnosis of epileptic seizure, and a comparative study of classifiers with classification accuracy have been made. For that, Hjorth's Descriptors, Recurrence Quantification Analysis, Shannon Entropy, Hurst Exponent, Zero Cross Rate methods, which have been used previously for diagnosis of disease or analysis of EEG signals, have been utilized. Obtained attributes have been firstly classified using k-nearest neighbor, naive Bayes, neural networks and support vector machines severally. For further performance improvement, the methods are evaluated in hybrid manner. For both detecting epileptic seizure, and distinguishing healthy and diseased people, 7 sub-data sets are created and accuracy, sensitivity and specificity of the methods were determined for the data sets. In this case, classification accuracy of between 97.5% and 100 % has been obtained and the results compared with the literature show that significant contributions have been made with this study.

Keywords: Epileptic Seizure, Hjorth, Entropy, Recurrence Plot, Hurst Exponent, Hybrid Classifier

ŞEKİLLER DİZİNİ

Şekil 1. Bir biyolojik sinir hücresinin yapısı	3
Şekil 2. Örnek EEG verileri.....	5
Şekil 3. Uluslararası 10/20 sistemine göre elektrotların kafa derisi üzerine yerleşimi	6
Şekil 4. Çalışmanın genel akış şeması.....	8
Şekil 5. İki farklı sinyalin tekrarlılık grafikleri	11
Şekil 6. Yapay sinir ağı modeli örneği	17
Şekil 7. Yapay sinir hücresinin yapısı	18
Şekil 8. Yapay sinir ağında katmanlar arası ilişki	19
Şekil 9. Tek katman ve iki girişe sahip algılayıcının yapısı	20
Şekil 10. 2 girişli algılayıcı için doğrusal ayrışma	20
Şekil 11. 3 girişli algılayıcı için doğrusal ayrışma	21
Şekil 12. Çok katmanlı yapay sinir ağı mimarisi	22
Şekil 13. k-EYK algoritması örneği	27
Şekil 14. Naive Bayes sınıflandırıcının graf yapısı	31
Şekil 15. İki sınıflı problemler için muhtemel hiper-düzlemler	32
Şekil 16. Optimum hiper-düzlem ve destek vektörleri.....	33
Şekil 17. Doğrusal ayrıştırılabilen veriler için hiper-düzlem ve destek vektörleri (sert marjin)	34
Şekil 18. Doğrusal ayrılabilen veriler için yumuşak marjin.....	36
Şekil 19. Doğrusal ayrıştırılamayan bir sınıflandırma problemi	37
Şekil 20. A,B,C,D ve E gruplarına ait örnek EEG verileri.....	41
Şekil 21. Normal ve epileptik nöbet dönemlerinin Hjorth betimleyicilerine göre uzayda dağılımları.....	43
Şekil 22. Normal, nöbet öncesi ve nöbet dönemlerinin Hjorth betimleyicilerine göre uzayda dağılımları	43
Şekil 23. Sağlıklı döneme ait tekrarlılık grafiği	44
Şekil 24. Nöbet öncesi döneme ait tekrarlılık grafiği.....	45
Şekil 25. Nöbet dönemine ait tekrarlılık grafiği.....	45
Şekil 26. Normal, nöbet öncesi ve epileptik dönemlere ait SGO dağılımları	46
Şekil 27. Normal, nöbet öncesi ve epileptik dönemlere ait Hurst üsteli dağılımları.....	47

Şekil 28. Normal, nöbet öncesi ve nöbet dönemine ilişkin entropi dağılımları	48
Şekil 29. Çapraz doğrulama işleminin yapısı	49
Şekil 30. Rastgele alt örnekleme (Random subsampling).....	49
Şekil 31. K-kat doğrulama.....	50
Şekil 32. ‘Birini Dışarıda Bırak’ doğrulama	50

TABLolar DİZİNİ

Tablo 1. Bayes için örnek veri seti	29
Tablo 2. Naive Bayes uygulaması için örnek veri kümesi	30
Tablo 3. Destek vektör makinelerinde kullanılan temel çekirdek fonksiyonları ve parametreleri.....	38
Tablo 4. Üzerinde çalışılacak olan veri grupları	41
Tablo 5. Kullanılan veri setleri ve açıklamaları	51
Tablo 6. Sınıflandırıcıya sunulacak olan öznitelik vektörleri ve boyutları	52
Tablo 7. 1. öznitelik vektörüne ilişkin sınıflandırma performansları	53
Tablo 8. Borda sayısı için örnek öncelik tablosu	55
Tablo 9. Hjorth betimleyicilerine göre oluşturulmuş öncelik tablosu.....	56
Tablo 10. Tüm alt kümeler için maksimum sınıflandırma doğrulukları	57
Tablo 11. Her bir öznitelik vektörü için maksimum sınıflandırma.....	58
Tablo 12. Her bir alt küme için ortalama sınıflandırma doğrulukları	58
Tablo 13. Her bir öznitelik vektörünün ortalama sınıflandırıcı.....	59
Tablo 14. Kullanılan melez sınıflandırıcı sistemleri	60
Tablo 15. Melez sınıflandırıcı sistemlerine ilişkin sınıflandırma performansları	60
Tablo 16. Sınıflandırma sonuçlarına ilişkin duyarlılık, özgünlük ve doğruluk oranları	61
Tablo 17. Sağlıklı, nöbet öncesi ve nöbet dönemlerine ilişkin performans tablosu.....	62
Tablo 18. Sağlıklı insanda göz açık durumda kaydedilen veriler ile nöbet öncesi ve nöbet durumlarının sınıflandırma performansları.....	62
Tablo 19. Nöbet verisinin diğer durumlardan ayırt edilme performansı.....	63
Tablo 20. Sağlıklı ve epilepsi hastası bireylerin genel olarak ayırt edilme performansları.....	64
Tablo 21. Sağlıklı ve nöbet olmayan durumdaki verilerin ayırt edilme performansları	64

SEMBOLLER DİZİNİ

AP	Aksiyon potansiyeli
EEG	Elektroensefalogram
ECoG	Elektrokortigram
TBA	Temel Bileşenler Analizi
EMG	Elektromiyogram
EOG	Elektrokülogram
HB	Hjorth Betimleyicileri
TG	Tekrarlılık Grafiği
TÖA	Tekrarlılık Ölçme Analizi
SGO	Sıfır Geçiş Oranı
HÜK	Hurst Üstel Katsayısı
ÖE	Örnek Entropi
YSA	Yapay Sinir Ağı
ÇKYSA	Çok Katmanlı Yapay Sinir Ağı
k-EYK	k-En Yakın Komşuluk
DVM	Destek Vektör Makinaları
BS	Bayes Sınıflandırıcı

1. GENEL BİLGİLER

1.1. Giriş

Bilgisayarlı sistemlerle hastalıkların tanısının gerçekleştirilmesi, son yıllarda en popüler konulardan biri haline gelmiştir. Bu sayede hastalıkların tanısı için geçen süre ve yanlış tanımlar azalmaktadır. Bir başka deyişle, gözle farkına varılamayacak durumda olan ayrıntıların yakalanması sağlanır. Bu teknoloji, epilepsi hastalığının teşhisinde de önemli rol oynamaktadır. Bu sebeple, otomatik tanı sistemlerinden faydalanılmaktadır. Bunun içinse, elektroensefalogram (EEG) ve elektrokortikogram (ECoG) gibi elektriksel işaretler kullanılır [1].

Epilepsi hastalığı, yaygın görülen tıbbi ve sosyal bir beyin fonksiyon bozukluğu hastalığıdır [2]. Tekrarlayan nöbetler şeklinde seyreden bu nörolojik hastalık, dünya nüfusunun yaklaşık %1'ini etkilemektedir. Her yaşta görülebilme olasılığı olan bu hastalık, insan yaşamını olumsuz şekilde etkilemektedir. Çünkü epilepsi hastalarında yetişkinlerde depresyon, anksiyete bozuklukları, psikoz; çocuklarda ise dikkat eksikliği, hiperaktivite, kekemelik gibi durumlara rastlanabilmektedir [2]. Bu nedenle biyomedikal araştırmalarda epilepsi teşhisi her geçen gün büyük önem kazanmaktadır.

Literatürde epilepsi işaretinin ve epileptik nöbetin teşhisine ve EEG veya ECoG işaretlerinin farklı amaçlar için sınıflandırılmasına ilişkin birçok çalışma yapılmış ve çeşitli öznelik çıkarma ve sınıflandırma yöntemleri önerilmiştir. Öznelik çıkarma yöntemlerinin zaman, frekans ya da zaman-frekans uzayında uygulanıp, uygun sınıflandırıcı ile sınıflandırılması şeklinde birçok çalışma mevcuttur. Srinivasan [3], EEG işaretlerinin uzaysal çözünürlüğünü iyileştirmek için yüzey Laplacian tekniğinden faydalanmıştır. Subasi ve Erçelebi tarafından gerçekleştirilen çalışmada [4], dalgacık dönüşümü kullanılarak öznelik çıkarımı ve yapay sinir ağı (YSA) ve lojistik regresyon kullanılarak sınıflandırma işlemi yapılmıştır. Ansari-Asl vd. çalışmalarında [5] senkronizasyon olasılığı ve fraktal boyutu öznelikleri kullanılarak doğrusal ayırtaç yöntemi ile EEG işaretleri sınıflandırılmıştır. Cecchin vd. [6] çalışmalarında işaretin ikinci dereceden ortalamasını ve baskın frekansını bulmak amacıyla iki Hjorth parametresini kullanmışlar ve %96 doğruluk oranı elde etmişlerdir. Orhan vd. [7], EEG işaretlerinin alt bantlarının sınıflandırılmasında dalgacık dönüşümünü, k-ortalama kümeleme algoritmasını

ve çok katmanlı yapay sinir ağını (ÇKYSA) kullanarak %95.60 civarında tanıma doğruluğu elde etmişlerdir. [3], [4], [5], [6] ve [7] çalışmaları, temel olarak EEG işaretinin alt bantlarının ayrıştırılmasına ya da farklı beyin bölgelerinden belli amaçlarla alınan EEG işaretlerinin birbirinden ayırt edilmesine ilişkin çalışmalardır. Epileptik nöbetin belirlenmesine yönelik çalışmalardan biri olan [8] çalışmasında, hasta ve sağlıklı insanlardan alınan normal döneme ve nöbet dönemine ait EEG işaretlerinin sınıflandırılması yapılmıştır. Entropi ve ayırık dalgacık dönüşümü yöntemleri kullanılarak, %96 tanıma doğruluğu elde edilmiştir. Subasi tarafından yapılan [9] çalışmasında ise epileptik nöbeti belirlemede dalgacık dönüşümü ve bulanık sinir ağından faydalanılmış ve %93.4 sistem duyarlılığı elde edilmiştir.

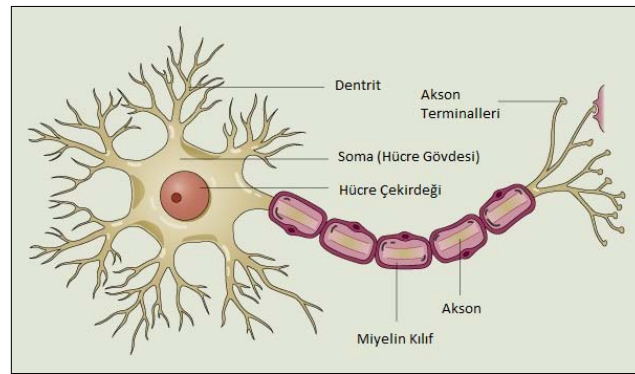
Almanya Bonn Üniversitesi, Epileptoloji Bölümü'nde Andrzejak vd. tarafından epilepsi teşhisi ve epileptik nöbet belirleme araştırmalarında kullanılmak üzere kaydedilen [10] ve bu tez çalışmasında da kullanılan EEG verileri ile de çeşitli çalışmalar mevcuttur. Gautama vd., doğrusal olmayan analiz yöntemleri olan tahmin hatası ve ilişki boyutu niteliklerinden faydalanarak, normal ve epileptik EEG işaretlerini karakterize etmişler ve %86.2 sınıflandırma başarıları sağlamışlardır [11]. Kannathal vd. ilişki boyutu, en büyük Lyapunov üssü, Hurst üstel katsayısı ve entropi ölçütlerini epileptik nöbetin tespitinde kullanmışlardır [12]. Yine başka bir çalışmada Kannathal vd. entropi özelliğinden ve adaptif bulanık mantık çıkarım sistemi (ABMÇS, ANFIS) sınıflandırıcıdan faydalanarak normal ve epileptik sinyallerin sınıflandırılmasında %90 başarı elde etmişlerdir [13]. Adeli vd. [14], epileptik nöbetin tespitinde dalgacık kaosu yöntemine başvurmuşlardır. Sadati vd. [15], çalışmalarında normal ve epileptik bölgelerin bölütlenmesi amacıyla uyarlanabilir nöral bulanık ağlar yöntemini öne sürmüşlerdir. Elde ettikleri sonuçları destek vektör makineleri (DVM), ANFIS ve ÇKYSA ile karşılaştırdıkları durumda, ANFIS ile yaklaşık %85.9 sınıflandırma doğruluğu elde etmişlerdir. Srinivasan vd., yaklaşık entropi yöntemini kullanan YSA tabanlı otomatik epileptik EEG belirleme sistemini öne sürmüşlerdir. Elman ve olasılıksal YSA yöntemlerinin kullanıldığı bu çalışmada %100 başarı vardır [16]. Tzallas vd., çalışmalarında zaman-frekans analizi metotlarına başvurmuş ve her bir segment için çıkarılan öznelikler k-en yakın komşuluk (k-EYK) sınıflandırıcı ile sınıflandırılarak %97.71 ila %100 arasında başarı elde etmişlerdir [17]. Polat ve Güneş, karar ağacı ve hızlı Fourier dönüşümü tabanlı melez bir sistem öne sürerek, epileptik nöbetin tespitini %98 doğrulukla gerçekleştirmişlerdir [18]. Martinez-Vargas vd. zaman-

frekans analizi ile öznitelik çıkarımı ve eşikleme fonkiyonu ile sınıflandırma yaparak %99 sınıflandırma başarısı elde etmişlerdir [19]. [20] çalışmasında, dalgacık dönüşümü ve çizgi uzunluğu yöntemleri ile YSA kullanılmış ve %97.77 başarı sağlanmıştır. [21] çalışmasında Dastidar vd., TBA ve YSA kullanarak %99.3 sınıflandırma başarısı elde etmişlerdir. Son olarak Liang vd. [22], zaman-frekans analizi ve entropi ölçütleri yardımıyla %98.67 sınıflandırma başarısı sağlamışlardır.

Yapılan çalışmanın adımlarına geçmeden önce beynin çalışmasına, otomatik tanı sistemine, EEG işaretinin karakteristiğine ve epilepsi hastalığına değinilecektir.

1.2. Beyin Hücresi ve Çalışması

Beyin, vücudun kontrol merkezidir. Tüm zihinsel faaliyetlerimizin kontrolünü beyin gerçekleştirir. Sinir sistemi, bilgi iletimini sağlayan nöron denilen hücrelerden oluşmaktadır [23]. Beynimizde yaklaşık 100 milyar nöron mevcuttur. Bu hücreler, elektrokimyasal etkileşimler sayesinde birbirleriyle haberleşirler. Her bir nöron 15.000 nöron ile bağlantı kurabilir ve her bir nöron diğer bir nörona 10 milisaniyeden daha kısa bir zamanda ulaşabilir. Bu süre göz kırpmaya süremizin onda birinden daha az bir süredir [24]. Bir nöronun yapısı Şekil 1'de [23] gösterilmiştir.



Şekil 1. Bir biyolojik sinir hücresinin yapısı

Bir nöron 3 temel kısımdan meydana gelir. Bunlar: dendrit, soma ve aksondur. Her nöron bu yapılar sayesinde, başka nöronlarla iletişim halindedir. Sinir sisteminde, nöronlar üzerinde taşınan elektriksel sinyale Aksiyon Potansiyeli (AP) denir [25]. Her nöron bu

sinyalleri kendisinden önceki nöronun dendrit denilen kısımlarla alır. Dendritler yardımıyla alınan bu sinyaller, soma denilen bölgede toplanarak işlenir. AP, akson boyunca ilerler ve akson ucunda bulunan telodendritler adı verilen dallanmış bölgeye ulaşır. Burada sinaps adı verilen boşluğa nörotransmitter maddelere salınarak, bir sonraki nöron yapısına bilgi iletimi gerçekleştirilir. Böylelikle, tek bir AP'nin nöronun nörona elektrobiyokimyasal şekilde iletimi gerçekleşmiş olur [25]. Bütün sinir sistemi, bu akışa göre çalışmaktadır.

Buradan görüldüğü üzere, beynimiz çok düşük şiddette sürekli olarak elektrik akımı üretir ve sinyalleri düzenli bir şekilde yayar. Bu sinyallerin bilgisayar ortamına kaydedilmesi işlemi ile EEG ya da ECoG işaretleri elde edilir ve otomatik tanı sistemlerinde kullanılırlar.

1.3. Otomatik Tanı Sistemleri

Otomatik tanı sistemleri, tıbbi görüntü ve işaretlerin otomatik olarak yorumlamasını sağlayan bilgisayar tabanlı sistemlerdir. Kanser türlerinin ve kalp ve beyin hastalıkları gibi hastalıkların tedavisinde kullanılırlar [26]. Bu sistemlerin kullanıldığı hastalıklardan biri de, yaygın bir beyin hastalığı olan epilepsi hastalığıdır.

Beynimizi oluşturan çok sayıdaki sinir hücresi (nöronlar), elektrobiyokimyasal etkileşimler ile birbirleriyle haberleşirler. Bu haberleşme sırasında, kafatası üzerine yerleştirilecek elektrotlar sayesinde, meydana gelen etkileşim hakkında bilgi edinmek mümkündür. Basitçe, beyin farklı işlemleri gerçekleştiren bölümlerden oluştuğu düşünüldüğünde, ilgili bölüme yakın yere yerleştirilecek olan elektrotlar, o bölgeye ilişkin bilgi edinilmesini sağlarlar.

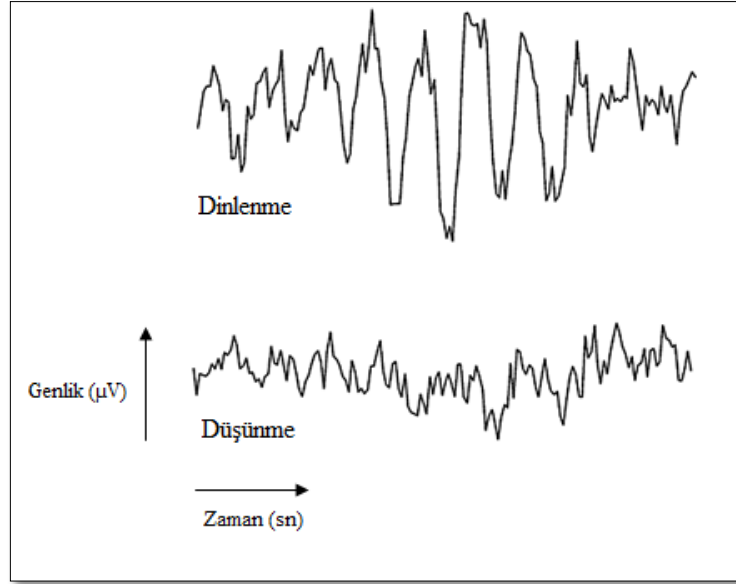
Otomatik tanı sistemlerinde, çeşitli tedaviler amacıyla elektroensefalogram (EEG) işaretleri kullanılmaktadır. EEG işaretleri, saç derisi üzerine yerleştirilen elektrotlarla kaydedilir. Bu nedenle pratiktir ve uygulaması kolaydır [27].

EEG kaydının temel amacı, sinir sisteminin çalışmasının, hangi beyin bölgesinin ne türden elektriksel bilgi yaydığına görülmesi ve takip edilmesidir. EEG, beyin lezyonlarının, tümörlerinin, enfeksiyonlarının tespiti; epileptik aktivitelerin, beyin ölümlerinin ve beyin elektriksel aktivitesini engelleyen her türlü hastalığın tanısında kullanılabilir [28].

Epilepsi hastalığının teşhisinde de EEG işaretleri üzerinden hastalık teşhisi konulmaktadır. Bir sonraki bölümde EEG işaretinin yapısına değinilecektir.

1.4. EEG İşaretlerinin Özellikleri ve Ölçülmesi

EEG işaretleri, periyodik ve durağan olmayan sinyaller olup; genlik, faz ve frekans değerleri sürekli değişmektedir. Bu nedenle yorumlanması oldukça zordur [29]. Dinlenme ve düşünme anındaki EEG verilerinin günümüz sistemleri yardımıyla kaydedilmiş halleri Şekil 2'de verilmiştir.



Şekil 2. Örnek EEG verileri

EEG işaretlerinin genlikleri 1-400 μV ve frekans bantları ise 0.5-100 Hz aralığındadır. Her ne kadar EEG sinyallerinin 0.5 ile 100 Hz arasında seyrettiği düşünülse de, klinik uygulamalarda ağırlık 0.5-30 Hz arasına verilmektedir. Bu frekans aralığında 4 frekans bandı mevcuttur [29]. Bunlar:

Delta Dalgaları: Derin uyku gibi beynin çok düşük aktivite gösterdiği durumlarda elde edilen; frekansları 0.5-4 Hz, genlikleri 20-400 μV arasında değişen dalgalardır.

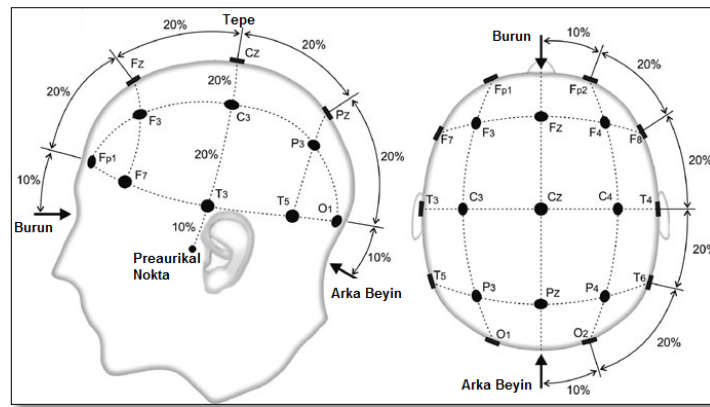
Teta Dalgaları: Rüya, stres gibi beynin düşük aktivite gösterdiği zamanlarda karşılaşılmaktadır. Frekansları 4-8 Hz, genlikleri 5-100 μV aralığında değişen dalgalardır

Alfa Dalgaları: Gözlerin kapalı ve fiziksel ve zihinsel dinlenme durumunda görülen dalgalardır. Frekansları 8-13 Hz, genlikleri 2-10 μ V aralığında değişir.

Beta Dalgaları: En yüksek aktivite düzeyine karşılık gelen dalgalardır. Frekansları 13 Hz'den fazla, genlikleri 1-5 μ V aralığında değişen dalgalardır [30].

EEG işaretinin düşük genlikli olmasının yanı sıra, beyin yapısının karmaşık olması ve karmaşık işlemleri kontrol etmesi ve elektromiyogram (EMG) ve elektrokülogram (EOG) gibi artifaktlardan ve ortamdaki elektromanyetik dalgalardan kolayca etkileniyor olması, bilgilere doğru şekilde ulaşılmasını zor hale getirmektedir. Çünkü bu tür durumlar, EEG işaretinin karakteristiğine zarar vermektedir. Bu sebeple, EEG işaretlerinin kaydedilmesi sırasında bahsi geçen bu artifaktlar göz önünde bulundurulmalı ve mümkün oldukça en aza indirilmelidir [29]. Bu sayede, EEG işaretlerinin yorumlanmasında daha sağlıklı sonuçlar elde edilebilecektir.

EEG işaretlerinin kaydedilmesinde, Uluslararası EEG ve Klinik Nörofizyoloji Federasyonu Birliği (The International Federation of Societies for Electroencephalography and Clinical Neurophysiology) tarafından klinik EEG ölçümleri için tavsiye edilen elektrot yerleştirme sistemi olan "Uluslararası 10-20 sistemi" kullanılmaktadır [31]. Çalışmada kullanılan epilepsi veri setinin elde edilmesinde de bu sistemden faydalanılmıştır. Şekil 3'te 10-20 sistemine göre elektrot dağılımının yapısı verilmiştir [31].



Şekil 3. Uluslararası 10-20 sistemine göre elektrotların kafa derisi üzerine yerleşimi

"10-20" adı, elektrotların dağılımdan ileri gelmektedir. Elektrotlar %10-%20-%20-%20-%20-%10 yoğunluğunca yerleştirilmektedir ve elektrot pozisyonları simetriktr [31].

Bu yerleşme şekline göre baş, dört standart nokta ile temsil edilmiştir. Noktalardan birincisi "nasion", buruna; ikincisi "inion", kafanın arka bölgesine ve diğer ikisi sağ ve sol preaurikülelere karşılık gelmektedir. "Nasion" ve "inion" boyunca, elektrotlar %10-%20-%20-%20-%20-%10 yoğunluğunca yerleştirilmektedir. Diğer elektrotlar ise, bu elektrotlar ile daire oluşturacak biçimde yerleştirilmektedir. Kulak memesine de toprak elektrotun yerleştirilmesi ile 20 adet elektrot kafa derisine yerleştirilmiş olur [31]. Bu sayede EEG işaretinin ölçülmesi ve EEG tabanlı BBA sistemlerinin oluşturulması sağlanmaktadır.

1.5. Epilepsi Hastalığının Tanımı ve Özellikleri

Epilepsi, tekrarlı ve ani nöbetler şeklinde meydana gelen bir sinir sistemi hastalığıdır [32]. Halk dilinde 'Sara hastalığı' olarak da bilinen epilepsi, kısa süreli beyin fonksiyon bozukluğu sonucunda ortaya çıkar. Beyin hücrelerinde meydana gelen geçici anormal elektrik akımları sonucunda, tekrarlayan ve ani epilepsi nöbetleri oluşur. Bu tekrarlayan nöbetler, zihinsel ve fiziksel fonksiyonları olumsuz şekilde etkiler.

İnsan beyni kafa travması, ateş yükselmesi gibi çeşitli nedenlerle epilepsi nöbeti geçirme potansiyeline sahiptir. İnsanların yaklaşık % 5'inin yaşamlarının herhangi bir dönemlerinde tek epilepsi nöbeti geçirdiği bilinmektedir. Ancak tek nöbet geçiren hasta epileptik değildir.

Epilepsi dünya nüfusunun yaklaşık %1'inde görülmekte ve günümüzde her yıl yaklaşık 30.000 yeni epilepsi hastasının topluma katıldığı tahmin edilmektedir [33]. Epilepsi hastalığı aslında memeli türlerinde görülür, bunun sebebi de memelilerin karmaşık bir beyin yapısına sahip olmalarıdır [2]. Yaygın bir nörolojik hastalık olan epilepsinin, ülkemizde de yaklaşık 720.000 bireyi etkilediği bilinmektedir. Bu hastaların %30'unu 18 yaş altı çocuklar oluşturmaktadır [34].

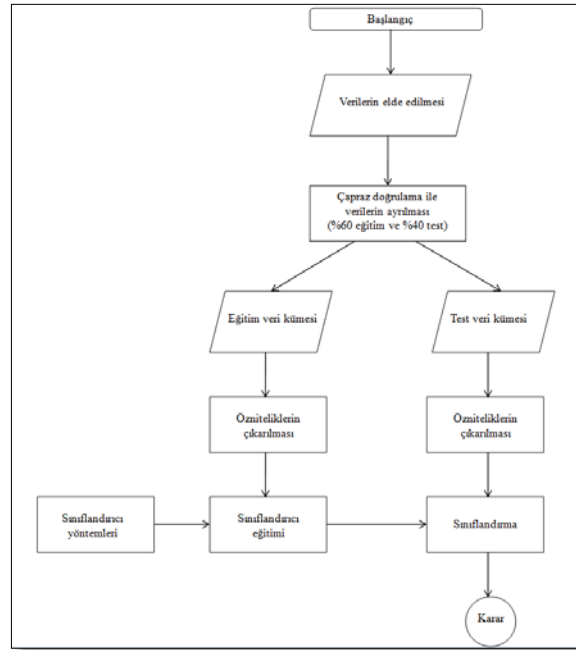
Epilepsi hastalığının sebepleri kesin olarak bilinmemekle beraber; beynin zarar görmesi veya yaralanması, kimyasal veya hormonal düzensizlikler, doğumda yaşanan problemler, kalıtım, ağır enfeksiyonlar ve tümörlerin epilepsi hastalığına sebep olduğu görülmüştür. Epilepsi; yaş, ırk, cinsiyet ve bunlar gibi hiçbir faktöre bağlı olmaksızın

herkeste görülebilir. Ancak, çocuklarda ve yaşlılarda ortaya çıkma ihtimalinin daha fazla olduğu bilinmektedir.

Literatürde 40'ın üzerinde nöbet tipi tanımlanmıştır [33]. Herkesin geçirdiği nöbet kendine özgü bazı farklılıklar barındırabilir. Bu nedenle bazı hastalarda epilepsi tanısının konulması zorlaşır ve problemin kaynağını bulmak uzun yıllar alabilir. Diğer hastalarda ise başka bir bozukluktan kaynaklanan belirtiler epilepsi olarak tanımlanabilir. Günümüzde, gelişen bilgisayar teknolojisiyle otomatik tanı yöntemleri yaygınlaşmış ve bu sayede yanlış tanımlar ve tanımlar için geçen süre azalmıştır [34].

1.6. Tezin Genel Yapısı ve Amacı

Çalışmada epilepsi hastalığına yönelik kaydedilen EEG işaretlerinin zamansal öznitelikleri değerlendirilerek, normal, nöbet öncesi ve epileptik nöbet dönemlerini ayırt eden bir EEG tabanlı otomatik tanı sisteminin elde edilmesi amaçlanmaktadır. Bunun için öncelikle verilerden öznitelikler elde edilmekte, çapraz doğrulama ile %60'ı eğitim ve %40'ı test verisi olacak şekilde ayrılmaktadır. Sınıflandırma işlemi sonucunda, verinin epileptik olup olmaması durumu incelenmektedir. Şekil 4'te sistemin genel akış şeması verilmektedir.



Şekil 4. Çalışmanın genel akış şeması

Öznitelik çıkarımında birbirleriyle ilişkili olarak Hjorth Betimleyicileri (HB), Tekrarlama Ölçme Analizi (TÖA), Sıfır Geçiş Oranı (SGO), Örnek Entropi (ÖE) ve Hurst Üstel Katsayısı (HÜK) yöntemleri ele alınmıştır. Test işleminin gerçekleştirilmesinde ise Yapay Sinir Ağı (YSA), k-En Yakın Komşuluk (k-EYK), Bayes Sınıflandırıcı (BS) ve Destek Vektör Makineleri (DVM) sınıflandırıcılarının tekli ve melez değerlendirilmeleri yapılmıştır.

1.7. Öznitelik Çıkarımı

Öznitelik çıkarımı bir sinyalin karakteristiğinin en iyi şekilde temsil edilmesi bakımından en önemli adımdır. Bu aşamada zaman uzayında öznitelikler değerlendirilmektedir.

1.7.1. Zaman Uzayında Öznitelik Çıkarımı

1.7.1.1. Hjorth Betimleyicileri

Hjorth betimleyicileri bir sinyalin ortalama değeri ve türevi ile ilişkili olarak elde edilen etkinlik, hareketlilik ve karmaşıklık parametrelerinden oluşan bir betimleyiciler kümesidir. Bu betimleyiciler 1970 yılında Hjorth tarafından EEG sinyallerinin analizi için önerilmiştir [35]. Etkinlik değeri işaretin varyansına eşit olmaktadır ve zaman uzayında N örnekten oluşan bir X işaretinin etkinlik değeri (1) eşitliği kullanılarak hesaplanmaktadır.

$$A(X_i) = \sigma_{X_i}^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1} \quad (1)$$

Burada \bar{X} , X işaretinin ortalamasını göstermektedir. Hareketlilik parametresi, işaretin ortalama frekansının bir ölçütüdür ve (2) eşitliğinden hesaplanmaktadır.

$$M(X_i) = \frac{\sigma_{\dot{X}_i}}{\sigma_{X_i}} = \frac{\sqrt{\frac{\sum_{i=1}^N (\dot{X}_i - \bar{\dot{X}})^2}{N-1}}}{\sqrt{A(X_i)}} \quad (2)$$

Burada \dot{X} , işaretin birinci dereceden türevine karşılık gelmektedir. Karmaşıklık parametresi ise işaretin sinüs işaretinden sapmasının bir ölçüsü olup, (3) eşitliği kullanılarak elde edilmektedir.

$$C(X_i) = \frac{\sigma_{\dot{X}_i} / \sigma_{\dot{X}_i}}{\sigma_{\dot{X}_i} / \sigma_{X_i}} \quad (3)$$

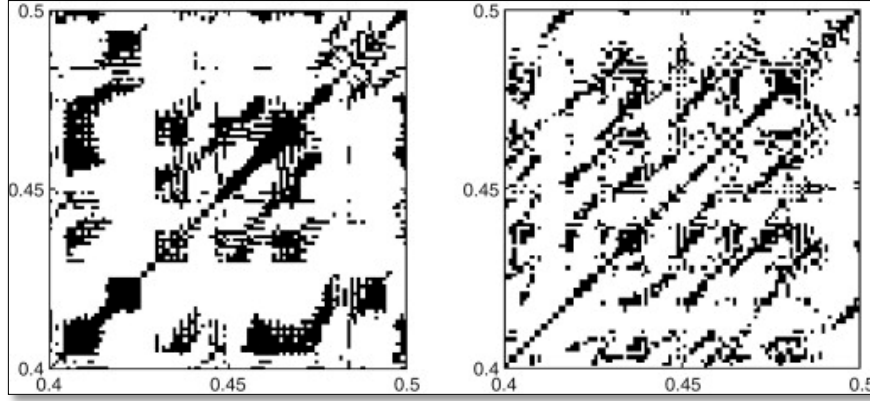
\ddot{X} , işaretin ikinci dereceden türevini ifade etmektedir [5] [36].

1.7.1.2. Tekrarlılık Grafiği (TG) ve Tekrarlılık Ölçme Analizi (TÖA)

1.7.1.2.1. Tekrarlılık Grafiği (TG)

Tekrarlılık grafiği, Eckmann [37] tarafından öne sürülen ve dinamik bir sistemde meydana gelen tekrarlı durumlara ortaya çıkaran bir matris ya da graftır. TG, zaman uzayında bir sinyalde kolaylıkla fark edilmeyen gizli periyodik durumların araştırılmasını sağlar. Bir zaman dizisi sinyalinin durağan olmama durumunu ölçmede faydalanılır [38]. TG yöntemi, doğrusal olmayan ayrıştırmada kullanılan avantajlı bir yöntemdir.

Tekrarlılık olduğundan bahsedilebilmesi için, i ve j gibi iki durum arasındaki uzaklık bir ϵ eşik değerinin altına düşmelidir. x_i , bir m boyutlu uzayda yörünge üzerinde i . nokta olsun. x_j , x_i 'ye yeterince yakın ise, grafik üzerindeki (i,j) konumuna bir nokta konur. x_i 'nin x_j 'ye olan uzaklığı, x_j 'nin x_i 'ye olan uzaklığına eşit olacağından, tekrarlılık grafiği $i=j$ eksenine göre simetriktir. Böylelikle, tekrarlılık grafiği $N \times N$ boyutta noktalardan oluşan bir matristir. Acharya vd. [39], çalışmalarında normal, nöbet öncesi ve nöbet dönemlerinin tanımlanmasında bu yöntemden faydalanmışlardır. Şekil 5'te iki farklı sinyale ilişkin tekrarlılık grafikleri verilmiştir.



Şekil 5. İki farklı sinyalin tekrarlılık grafikleri

1.7.1.2.2. Tekrarlılık Ölçme Analizi (TÖA)

Tekrarlılık grafiğinin yorumlanmasında tekrarlılık ölçme analizi (TÖA) parametrelerinden faydalanılır. TÖA parametreleri, bir zaman uzayı sinyalindeki durağan olmama ve gizli periyodiklik durumları analiz eder [37]. Acharya vd. [38], [39], epileptik nöbetin tespiti için EEG sinyallerinin sınıflandırılmasında TÖA parametrelerinden faydalanmıştır. $R_{i,j}$ tekrarlılık grafiği olmak üzere, kullanılan TÖA parametreleri aşağıdaki gibi hesaplanmaktadır:

Tekrarlılık Oranı (Recurrence Rate): Bir tekrarlılık grafiğindeki, tekrarlayan noktaların yüzdesidir. Korelasyon toplamına karşılık gelir.

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j} \quad (4)$$

Belirlilik (Determinism): Çapraz hatlar üzerine konumlanmış tekrarlayan noktaların yüzdesidir.

$$DET = \frac{\sum_{l=l_{\min}}^N IP(l)}{\sum_{i,j}^N R_{i,j}} \quad (5)$$

$P(l)$, l uzunluklu çapraz hatların histogramıdır.

Tabakalanma (Laminarity): Dikey hatlar üzerine konumlanmış tekrarlayan noktaların yüzdesidir.

$$LAM = \frac{\sum_{\nu=\nu_{\min}}^N \nu P(\nu)}{\sum_{\nu=1}^N \nu R(\nu)} \quad (6)$$

$P(\nu)$, ν uzunluklu dikey hatların histogramıdır.

Belirlilik ve Tekrarlılık Oranı İlişkisi: Belirlilik ve tabakalanma parametreleri arasındaki orandır.

$$RATIO = N^2 \frac{\sum_{l=l_{\min}}^N lP(l)}{\left(\sum_{l=1}^N lP(l)\right)^2} \quad (7)$$

Ortalama Çapraz Hat Uzunluğu (L): Çapraz hatların ortalama uzunluğudur.

$$L = \frac{\sum_{l=l_{\min}}^N lP(l)}{\sum_{l=l_{\min}}^N P(l)} \quad (8)$$

Yakalama Zamanı (Trapping Time): Dikey hatların ortalama uzunluğudur.

$$TT = \frac{\sum_{\nu=\nu_{\min}}^N \nu P(\nu)}{\sum_{\nu=\nu_{\min}}^N P(\nu)} \quad (9)$$

En Uzun Çapraz Hat Uzunluğu:

$$L_{\max} = \max(\{l_i; i = 1 \dots N_l\}) \quad (10)$$

En Uzun Dikey Hat Uzunluđu:

$$V_{\max} = \max(\{v_i; i = 1 \dots N_v\}) \quad (11)$$

Iraksaklık (Divergence): L_{\max} 'in tersidir.

$$DIV = \frac{1}{L_{\max}} \quad (12)$$

Düzensizlik (Entropy): Çapraz hatların uzunluklarının olasılıksal dağılımlarının Shannon entropisi değeridir.

$$ENTR = - \sum_{l=L_{\min}}^N p(l) \ln p(l) \quad (13)$$

1.7.1.3. Sıfır Geçiş Oranı

Sıfır geçiş oranı (SGO), bir işaret boyunca var olan işaret değışikliklerinin oranıdır. Başka bir deyişle, işaretin sıfır eksenini kaç kez keştiğinin bir ölçütüdür. Bu öznelik genellikle ses işaretlerinden vurmali sesleri tespit etmek amacıyla tercih edilir [40]. Ancak EEG ve ECoG gibi farklı frekans bileşenlerinden meydana gelen işaretler de karakteristik olarak ses işaretleriyle benzerlik gösterdiğinden, bu çalışmada sıfır geçiş oranından faydalanılmıştır ve literatürde bu kapsamda birçok çalışma mevcuttur. Örneğin [41] çalışmasında Alzheimer hastalığının teşhisinde EEG işaretinin zamansal özneliklerinden biri olarak sıfır geçiş oranı kullanılmıştır. SGO,(14) eşitliği kullanılarak hesaplanmaktadır.

$$ZCR = \frac{1}{2N} \sum_{n=1}^N |\text{sign}(x[n]) - \text{sign}(x[n-1])| \quad (14)$$

1.7.1.4. Hurst Üstel Katsayısı

Hurst üstel katsayısı, zaman serilerinin sınıflandırılması amacıyla kullanılan istatistiksel bir ölçüttür. Hurst üstel katsayısı, sürekli ve ayrık zaman uzayı işaretlerinin davranışı hakkında bilgi verir ve bu parametre epileptik nöbetin belirlenmesinde EEG ya da ECoG sinyalleri üzerinde kullanılır [42].

Hurst üstel katsayısını tanımlamanın çeşitli yolları vardır ancak bunlardan en eskisi Harold Hurst tarafından tanımlanan eşitliklerdir ve buna “Yeniden Ölçeklenmiş Genişlik (Rescaled Range)” de denir. R/S oranı buradan gelmektedir. (15) eşitliğindeki gibi ifade edilmektedir.

$$E \left[\frac{R(n)}{S(n)} \right] = cn^H, n \rightarrow \infty \quad (15)$$

X_i , ($i=1,2,\dots,n$) şeklinde tanımlanan bir zaman uzayı dizisi olmak üzere, R/S oranı aşağıdaki adımlarla hesaplanır:

İlk olarak serinin ortalaması (m) bulunur.

$$m = \frac{1}{n} \sum_{i=1}^n X_i \quad (16)$$

Ortalama değer, orjinal değerden çıkarılarak sapma serisi denilen Y_i dizisi elde edilir ve bu serinin kümülatif toplamı hesaplanır.

$$Y_t = X_t - m, (t = 1, 2, 3, \dots, n) \quad (17)$$

$$Z_t = \sum_{i=1}^t Y_i \quad (18)$$

Kümülatif sapma serisi aralığı (R) ve bu serinin standart sapması (S) hesaplanır.

$$R_t = \text{Max}(Z_1, Z_2, Z_3, \dots, Z_t) - \text{Min}(Z_1, Z_2, Z_3, \dots, Z_t) \quad (19)$$

$$S_t = \sqrt{\frac{1}{t} \sum_{i=1}^t (X_i - u)^2} \quad (20)$$

Yeniden ölçeklenmiş genişlik değeri (R/S) elde edilir.

$$(R/S)_t = R_t / S_t \quad (21)$$

Zaman periyodu uzadıkça R/S değerinin n. dereceden kuvveti alınır. Hurst üstel katsayısı, bu kuvvetin derecesidir.

$$(R/S)_t = cxt^H \quad (22)$$

Burada c, keyfi bir sabit; H ise Hurst üstel katsayısıdır. H'nin bulunabilmesi için her iki tarafın da logaritması alındığında, $\log(R/S) = \log c + H \log(t)$ ifadesi elde edilir. Bu eşitlik sayesinde R/S ile gözlem sayısı t arasında bir ilişki kurulur. Gözlem sayısının değeri, parametrenin güvenilirliğini etkileyen faktördür.

1.7.1.5. Entropi

Entropi, bir durumdaki ya da değişkendeki bilgi miktarını veren bir ölçüttür [43]. Başka bir deyişle, bir durumdaki ya da değişkendeki belirsizlik değeridir. Sinyal açısından bakıldığında bir sinyalin entropisi, o sinyalin içerdiği bilgi miktarı hakkında bir ölçüdür; bu da işaretin dağılımına karşılık gelir. Geniş, düz bir olasılıksal dağılıma sahip verinin entropisi yüksek değerdedir. Tersine şekilde dar, sivri dağılıma sahip olan verinin entropisi ise düşük değerdedir.

Entropi, EEG işaretleri içerisindeki değişkenliği gösteren bir başka istatistiksel tanımlayıcıdır. Sonlu uzunlukta sayısal bir rastgele değişkenin olasılık dağılım fonksiyonu $p(x)$ ile ifade edilirse, bu değişkenin entropisinin hesaplanması (23) eşitliğinde ifade edilmiştir. Aynı zamanda bu eşitlik Shannon entropisi olarak da adlandırılır [44].

$$H(x) = - \sum_{i=1}^{N-1} p_i(x) \log_2(p_i(x)) \quad (23)$$

x bir olasılıksal deney; $p(x)$ ise bu x sonucunun ortaya çıkma olasılığı ise, bu deneydeki ortalama bilgi miktarı $H(X)$ 'tir. Shannon entropi ölçütünden başka, yaklaşık entropi de uygulamalarda sıkça kullanılmaktadır.

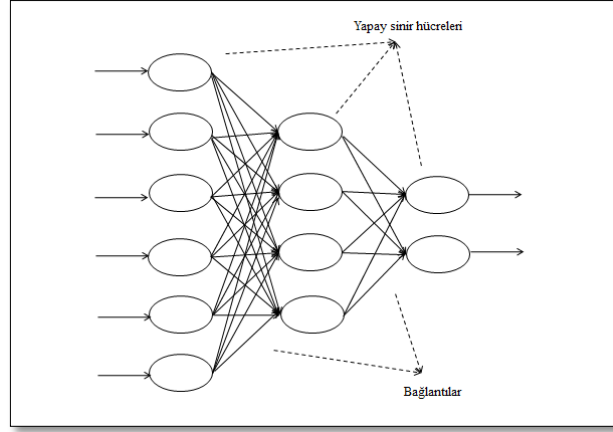
1.8. Sınıflandırma Yöntemleri

Sınıflandırma aşamasında, test verisi olarak ayrılan veri seti ile eğitim verisi olarak ayrılan veri setinin özelliklerinin karşılaştırılması işlemi gerçekleştirilmektedir. Bu adımda çok katmanlı yapay sinir ağları, k-en yakın komşuluk, Bayes sınıflandırıcı ve destek vektör makinelerinden yararlanılmıştır.

1.8.1. Yapay Sinir Ağları

Yapay sinir ağları; insan beyninin yerine getirebildiği öğrenme, yeni bilgiler türetme ve keşfetme gibi yeteneklerinin bilgisayar sistemlerine uyarlanmasını sağlayan, öğrenmeye dayalı bir yöntemdir. İnsan beyni olayları öğrenebilen ve bu olaylar üzerinden çıkarım yapabilen bir yeteneğe sahiptir. Yapay sinir ağlarının tasarımında; insan beyninin öğrenme, ilişki kurma, sınıflandırma, genelleme yapma, özellik belirleme ve optimizasyon gibi yetilerinin bilgisayar sistemleri tarafından da gerçekleştirilebileceği fikri ortaya atılmıştır [45].

YSA sistemleri, insandaki biyolojik sinir sisteminin temelini oluşturan sinir hücrelerinin yani nöronların çalışma mekanizmasından esinlenilerek geliştirilen bir sınıflandırma tekniğidir. Nöron adı verilen yapay sinir hücrelerini birbirine bağlayan bağlantılar vardır ve her biri bir değere sahiptir. YSA, yapay sinir hücrelerinin birbirleriyle çeşitli şekillerde bağlanmasıyla elde edilir ve katmanlar şeklinde tasarlanabilir [46]. Örnek bir YSA modeli Şekil 6'da verilmiştir.



Şekil 6. Yapay sinir ağı modeli örneği

Bir yapay sinir ağının görevi, kendisine verilen bir girdi setine göre bir çıktı seti belirlemektir. Bunun gerçekleşebilmesi için ağ, ilgili örneklerle eğitilerek, genelleme yapacak duruma getirilir. Bu genellemeler sayesinde benzer girdilere karşılık gelen çıktı setleri elde edilir [47].

Yapay sinir ağları bilinen hesaplama yöntemlerinden farklı bir yöntem sunmaktadır. Çalışma prensipleri ve özellikleri, kullanılan ağ modeline göre değişmektedir. Ancak her ağda geçerli olan birtakım özellikler mevcuttur. Bunlara değinilecek olursa, ilk olarak her yapay sinir ağının görevi bilgisayarların öğrenmesini sağlayarak, olayları öğrenme yolu ile karara ulaştıran bir sistem tasarlamaktır. Yapay sinir ağlarının olayları öğrenebilmeleri için, o olayla ilgili örneklerin belirlenmesi ve yeterli sayıda olmaları gerekmektedir. Aksi takdirde ağın eksik öğrenmesi ya da ezberle dayalı karar vermesi durumu oluşur. Bunun için de, ağın muhakkak test edilmesi gerekmektedir [47].

Yapay sinir ağlarının bir önemli özelliği de, görülmemiş örnekler hakkında dahi bilgi üretebilmeleridir. Bu durum, kendisine gösterilen örneklerden yola çıkarak genelleme yapma özellikleri sayesinde gerçekleşir. Aynı zamanda yapay sinir ağları, eksik bilgi ile de çalışabilirler. Bir ağ, eğitim aşamasını geçtikten sonra, eksik bilgi içeren bir girdi seti dahi gelse, verimli sonuç üretebilir. Bu da yapay sinir ağının hata toleransına açık olma özelliğini ortaya koyar. Böylece ağ yapısı eksik bilgidan dolayı bozulmamış, sadece ilgili nöron etkilenmiş olur [46].

Başka bir özellik ise, yapay sinir ağlarında bilginin saklanması biçimidir. Yapay sinir ağlarında bilgi, ağın bağlantı değerleri ile ölçülmekte ve yine ağın bağlantılarında

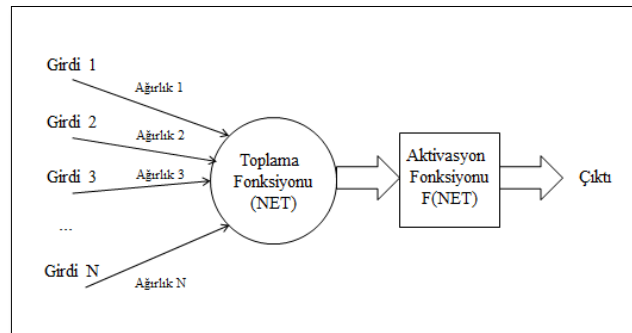
saklanmaktadır. Diğer sistem türlerinde olduğu gibi, bir veri tabanında saklama ya da program içine gömülme durumu yoktur. Bu nedenle, diğer yöntemlere göre yorumlanması daha zordur. Yapay sinir ağlarında, bir bağlantı tek başına bir anlam taşımaz. Bilgi ağa dağıtılmış durumdadır. Bu da ağın bilgilerinin açıklanamamasının bir sebebidir [46].

Yapay sinir ağlarının birçok alanda kullanımı mevcuttur. Ancak yukarıda bahsedilen özelliklerin yanı sıra, ağın kendisine has özellikleri de dikkate alındığında, yapay sinir ağları günümüzde özellikle örüntü tanıma, sinyal filtreleme, tıbbi yorumlamalar, veri madenciliği vb. alanlarda en çok kullanılan ve geliştirilen yöntemlerden biri olmuştur [48].

Yapay sinir ağları biyolojik sinir sisteminin çalışmasından yola çıkılarak geliştirilmiş bir modeldir. Kısaca biyolojik sinir ağlarına değinilecek olursa; biyolojik sinir ağları milyonlarca sinir hücresinden (nöron) meydana gelen bir sistemdir. Bu hücreler birbirlerine bağlanarak beyin fonksiyonlarını yerine getirirler. Bir biyolojik sinir hücresinin yapısı Şekil 1'de verilmişti. Beynimizde 1010 adet sinir hücresi ve bunların da birbirleriyle 6×10^{13} 'ten fazla sayıda bağlantısı olduğu varsayılmaktadır ve buna rağmen insan beyni oldukça hızlı çalışarak fonksiyonları yerine getirme özelliğine sahiptir. Biyolojik sinir hücresinin bu şekilde çalışması, çok hızlı çalışan mükemmel bir bilgisayar sistemi olarak düşünülmüş ve buradan yola çıkılarak yapay sinir sistemi fikri ortaya atılmıştır.

1.8.1.1. Yapay Sinir Hücresi (Nöron)

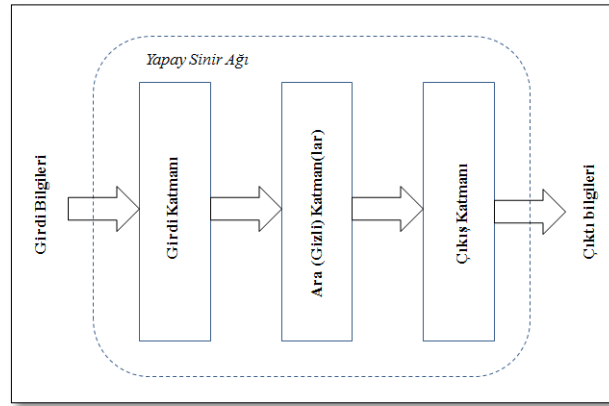
Yapay sinir ağları ile insan beyninde var olan yeteneğin bilgisayarlara kazandırılması amaçlanmıştır. Biyolojik sinir sistemlerinde olduğu gibi, yapay sinir sistemlerinin de yapay sinir hücreleri vardır. Bir yapay sinir hücresinin yapısı Şekil 7'de gösterilmiştir [47].



Şekil 7. Yapay sinir hücresinin yapısı

Şekil 7'den de görüldüğü gibi, her bir yapay sinir hücresi 5 temel kısımdan oluşur: girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıktı. Girdiler bir yapay sinir hücresine dış dünyadan, başka bir sinir hücresinden ya da kendi kendisinden gelen bilgilerdir. Bu girdiler, ağırlık değerleriyle anlamlandırılır. Toplama fonksiyonuna girdiler kendi ağırlıkları ile çarpılarak alınır ve burada hücreye gelen net girdi hesaplanır. Toplama fonksiyonunda hesaplanan net girdi aktivasyon fonksiyonu ile işlenerek, buna karşılık gelen çıktı değeri üretilir. Genel olarak bir yapay sinir hücresi bu şekilde çalışır ve birden fazla hücrenin bir araya gelmesiyle yapay sinir ağı meydana gelir.

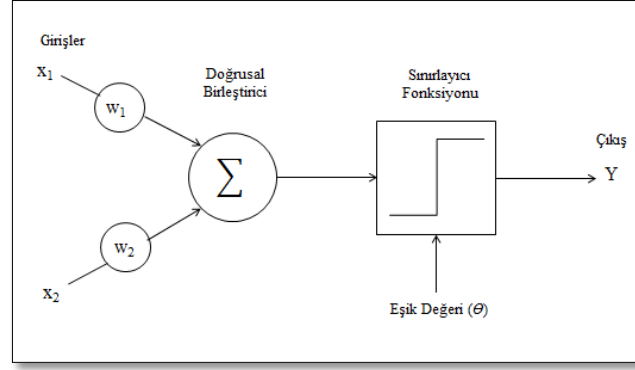
Bu hücrelerin bir araya gelmesinde bir sistematik vardır. Hücreler genel olarak 3 ana katman üzerinde ve her katman içinde paralel olacak şekilde oluşturulur. Bu katmanlar sırasıyla girdi katmanı, ara katman(lar) ve çıktı katmanıdır. Girdi katmanı bilgilerin ağı giriş yaptığı katmandır. Bu bilgiler ara katmanda işlenerek çıktı katmanına gönderilirler. Çıktı katmanı ile üretilen çıktı, karar olarak dışarıya gönderilir [47]. Şekil 9'da yapay sinir ağının katmanları arasındaki ilişki verilmiştir.



Şekil 8. Yapay sinir ağında katmanlar arası ilişki

1.8.1.2. Algılayıcı

Algılayıcı, yapay sinir ağlarının ilk ortaya atılan ve en basit yapısıdır. Şekil 9'da tek katman ve iki girişe sahip algılayıcının yapısı verilmiştir.

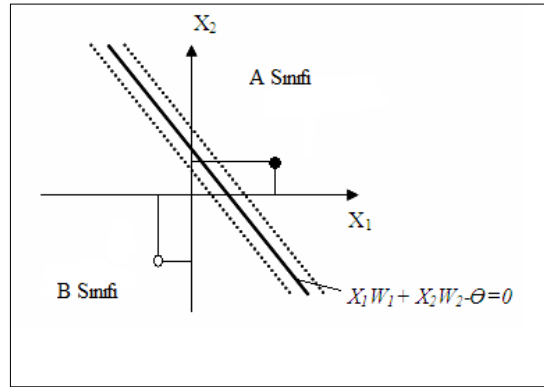


Şekil 9. Tek katman ve iki girişe sahip algılayıcının yapısı

Algılayıcının yapısında sırasıyla ağırlıklı girişler, doğrusal birleştirici ve bunu izleyen bir sınırlayıcı vardır. Girişlerin ağırlıklı toplamları sınırlayıcı fonksiyonuna verilerek, çıkışta 1 veya (-1) değerleri alınır. Çünkü algılayıcıda amaçlanan, girişleri A ve B gibi iki sınıfa yerleştirmektir. Böylece N boyutlu bir giriş uzayı, iki karar bölgesine ayrılmış olur. Hiper-düzlem fonksiyonu doğrusal ayrılabilir fonksiyonla (24) tanımlıdır.

$$\sum_{i=1}^n x_i w_i - \theta = 0 \quad (24)$$

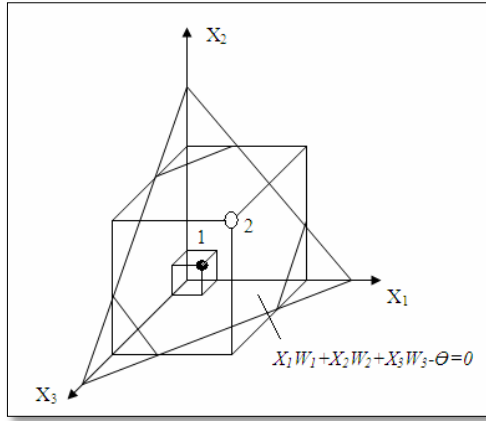
2 giriş olması durumunda x_1 ve x_2 karar sınırları oluşur. Karar sınırının ötelenmesinde θ değeri kullanılır. Şekil 10'da 2 girişli algılayıcı için doğrusal ayrışmaya bir örnek verilmiştir.



Şekil 10. 2 girişli algılayıcı için doğrusal ayrışma

3 giriş olması durumunda ise algılayıcı (25) eşitliğiyle ifade edilir. Bu durumda 3 giriş bir düzlemle tanımlanmış olur. Bu düzlem örneği ise Şekil 11'de gösterilmiştir.

$$x_1w_1 + x_2w_2 + x_3w_3 - \theta = 0 \quad (25)$$



Şekil 11. 3 girişli algılayıcı için doğrusal ayrışma

Algılayıcıda sınıflandırma işlemi, gerçek çıkış ile hedeflenen çıkış arasındaki farkı azaltacak şekilde ağırlıkların ayarlanması ile gerçekleştirilir. Başlangıç ağırlıkları rastgele olarak atanır ancak genellikle $[-0.5, 0.5]$ aralığındadır. Çıkış değerlerinin eğitim örnekleriyle tutarlı olması için ağırlıklar güncellenir. Algılayıcı için ağırlık güncelleme işlemi aşağıda verilmiştir.

$$e(p) = Y_h(p) - Y(p), \quad p = 1, 2, 3, \dots \quad (26)$$

Burada $Y(p)$, p . iterasyonda hesaplanan çıkış ve beklenen çıkış ise $Y_k(p)$ 'dir. p değeri, p . örneğin algılayıcıya verilmesini temsil eder. Eğer hata $e(p)$, pozitifse, $Y(p)$ çıkışı artırılmalı; negatifse azaltılmalıdır. Algılayıcı öğrenme kuralı (27)'den hesaplanır.

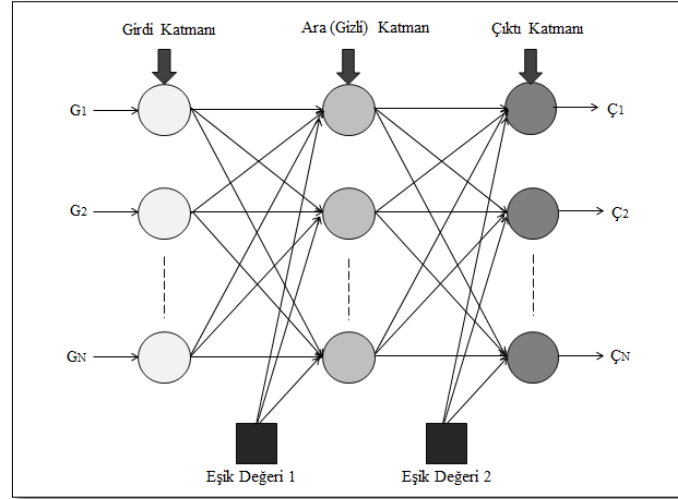
$$w_i(p+1) = w_i(p) + \alpha * x_i(p) * e(p) \quad (27)$$

Burada α , pozitif ve 1'den küçük bir sabit olup, "öğrenme katsayısı" olarak adlandırılır.

Algılayıcı, sadece doğrusal olarak ayrılabilen sınıflandırma problemlerini çözebildiğinden ve gerçek hayattaki problemler her zaman doğrusal olmayacağından bu amaç için yetersiz kalmaktadır. Bunun içinse, çok katmanlı yapay sinir ağları kullanılır.

1.8.1.3. Çok Katmanlı Yapay Sinir Ağları

Çok Katmanlı Yapay Sinir Ağları (ÇKYSA), bir ya da daha fazla katmandan oluşan sinir ağı modelidir. Girdi katmanından giriş verileri alınır ve ara katmandaki nöronlara dağıtılır. Çıktı katmanı ara katmandan gelen bilgiyi alarak, çıkış değerini üretir. Şekil 12'de ÇKYSA mimarisi verilmiştir.



Şekil 12. Çok katmanlı yapay sinir ağı mimarisi

Ara katman(lar)daki nöronlar, özellikleri tespit ederler. Bu sebeple ara katmandaki katman ve nöron sayısı, ayrışmadaki önemli parametrelerdir.

1.8.1.4. ÇKYSA'nın Öğrenme Kuralı

ÇKYSA'nın öğrenme kuralı en küçük kareler yöntemine dayalı Delta Öğrenme Kuralının geliştirilmiş hali olduğundan Genelleştirilmiş Delta Kuralı olarak adlandırılır ve iki safhadan oluşur: ileri doğru hesaplama ve geriye doğru hesaplama. İleri doğru hesaplamada ağın çıkışı hesaplanır, geriye doğru hesaplamada ise beklenen çıkışa göre

ağın ağırlık değerleri yeniden hesaplanarak güncellenir. Böylelikle kabul edilir bir hata düzeyine gelene dek ağ eğitilmiş ve sağlıklı bir sistem elde edilmiş olur.

1.8.1.4.1. İleri Doğru Hesaplama

Eğitim veri setinden bir örnek, girdi katmanı üzerinden ağa verilir ve burada herhangi bir işleme uğramadan ara katmana gönderilir.

$$C_k^i = G_k \quad (28)$$

Ara katman(lar)a gelen bilgiler bağlantı ağırlıklarıyla (A_1, A_2, \dots) birlikte katmana ulaştığından dolayı, ara katmandaki nöronlara gelen NET giriş (29) eşitliğindeki gibi hesaplanır.

$$NET_j^a = \sum_{k=1}^n A_{kj} C_k^i \quad (29)$$

A_{kj} , k. girdi elemanını j. ara katman elemanının bağlayan bağlantıya ait ağırlık değerini göstermektedir. j. ara katmanın çıkışı da, bu net girişin aktivasyon fonksiyonunun geçirilmesi ile hesaplanır. Geriye doğru hesaplama yapılacağından, aktivasyon fonksiyonunun türevi alınabilen bir fonksiyon olması gerekir ve genellikle sigmoid fonksiyonu kullanılır. Bu durumda çıkış aşağıdaki gibi olur.

$$C_j^a = \frac{1}{1 + e^{-(NET_j^a + \beta_j^a)}} \quad (30)$$

Burada β_j , ara katmandaki j. elemana bağlanan eşik değer elemanının ağırlığını göstermektedir. Bu eşik değerinin birimi çıkışı sabittir ve 1'dir. Ağırlık değeri ise, sigmoid fonksiyonunun uyumunu belirlemek amacıyla konmuştur. Eğitim esnasında bu değer ağ tarafından belirlenmektedir.

Ara katmanın ve çıkış katmanının her bir elemanının çıkışları aynı şekilde NET girişlerin hesaplanması ve aktivasyon fonksiyonundan geçirilmesi ile elde edilir. Çıktı katmanındaki çıkışlar (C_1, C_2, \dots) bulununca ileri hesaplama tamamlanır.

1.8.1.4.2. Geriye Doğru Hesaplama

Ağa verilen giriş için üretilen çıkışlar, beklenen çıkışlarla (B_1, B_2, \dots) karşılaştırılır. Aradaki fark hata değerini verir. Geriye doğru hesaplamada, bu hata ağırlık değerlerine dağıtılıp, bir sonraki iterasyonda bu hatanın azaltılması sağlanır. Çıktı katmanında m . eleman için oluşan hata değeri (31) eşitliğinde verilmiştir.

$$E_m = B_m - C_m \quad (31)$$

Burada E_m yalnızca bir eleman için oluşan hata değeridir. Çıkış katmanında oluşan toplam hatayı (TH) bulmak için, her bir elemana ait hataların toplanması gerekir. Hata değerlerinin negatif olabilme durumundan dolayı, bu hataların kareleri hesaplanarak toplanır ve sonucun karekökü alınır. ÇKYSA'da amaç, bu hatayı en aza indirecek şekilde ağı eğitmektir. TH'nin hesaplanması (32)'de verilmiştir.

$$TH = \frac{1}{2} \sum_m E_m^2 \quad (32)$$

Toplam hatanın en aza indirilmesi için, bu hataya neden olan elemanlara hatanın dağıtılması yani mevcut elemanların ağırlıklarının değiştirilmesi gerekir. Değiştirilecek ağırlıklar "ara katman-çıkış katmanı" veya "ara katman-girdi katmanı" arasındaki ağırlıklardır.

Ara katmandaki j . elemanı çıktı katmanında m . elemana bağlayan bağlantıdaki değişim miktarı ΔA_a ise, t . iterasyonda ağırlık değişim miktarı (33)'teki gibi hesaplanır.

$$\Delta A_{jm}^a(t) = \lambda \delta_m C_j^a + \alpha \Delta A_{jm}^a(t-1) \quad (33)$$

λ öğrenme katsayısı, α ise momentum katsayısıdır. Öğrenme katsayısı ağırlıkların değişim miktarını, momentum katsayısı ise ÇKYSA'nın öğrenme aşamasında yerel minimum ya da maksimum noktaya takılıp ağırlıkların aksamasına engel olmak adına ağırlık değişim değerinin belli bir oranda bir sonraki değişime eklenmesini sağlar. δ_m , m. çıkış elemanının hatasını göstermektedir ve (34) eşitliğindeki gibi hesaplanır.

$$\delta_m = f'(NET) \cdot E_m \quad (34)$$

$f'(NET)$ aktivasyon fonksiyonun türevine karşılık gelir ve sigmoid fonksiyon kullanılması durumunda eşitlik (35)'teki gibi verilir.

$$\delta_m = \zeta_m (1 - \zeta_m) E_m \quad (35)$$

Böylelikle ağırlıkların t. iterasyondaki değerleri (36)'dan elde edilir.

$$A_{jm}^a(t) = A_{jm}^a(t-1) + \Delta A_{jm}^a(t) \quad (36)$$

Çıktı katmanında bulunan elemanların eşik değer ağırlıklarının (β^ζ) değişim miktarları, bu birime ait çıkışın 1 olması ile (37)'deki gibi hesaplanır.

$$\Delta \beta_m^\zeta(t) = \lambda \delta_m + \alpha \Delta \beta_m^\zeta(t-1) \quad (37)$$

Eşik değerinin t. iterasyon sonucu ağırlığın yeni değeri ise (38)'de verilmektedir.

$$\beta_m^\zeta(t) = \beta_m^\zeta(t-1) + \Delta \beta_m^\zeta(t) \quad (38)$$

Çıktı katmanında meydana gelen hatalarda, sadece ara katman ve çıktı katmanının değil, aynı zamanda girdi katmanı ve ara katman arasındaki ağırlıkların da etkisi olacağından bu iki katman arasındaki ağırlıkların da güncellenmesi gerekmektedir. Bunun için çıktı katmanındaki elemanların hepsinin hatası işleme alınır. Ara katman ve çıktı

katmanındaki benzer şekilde, ağırlıklardaki değişim miktarı ΔA_i , hata terimi δ^α , ağırlıkların yeni değerleri A_{kj}^i ve eşik değer ağırlıkları β_j^a (39), (40), (41), (42), (43) ve (44) eşitliklerinden elde edilir.

$$\Delta A_{kj}^i(t) = \lambda \delta_j^a \zeta_k^i + a \Delta A_{kj}^i(t-1) \quad (39)$$

$$\delta_j^a = f'(NET) \cdot \sum_m \delta_m A_{jm}^a \quad (40)$$

$$\delta_j^a = \zeta_j^a (1 - \zeta_j^a) \cdot \sum_m \delta_m A_{jm}^a \quad (41)$$

$$A_{kj}^i(t) = A_{kj}^i(t-1) + \Delta A_{kj}^i(t) \quad (42)$$

$$\Delta \beta_j^a(t) = \lambda \delta_j^a + a \Delta \beta_j^a(t-1) \quad (43)$$

$$\beta_j^a(t) = \beta_j^a(t-1) + \Delta \beta_j^a(t) \quad (44)$$

Bir iterasyon ileri ve geri hesaplanmaların yapılmasıyla tamamlanır ve öğrenme gerçekleşene kadar bu işlem her bir örnek için tekrarlanır.

1.8.1.5. ÇKYSA'nın Çalışma Prensipleri

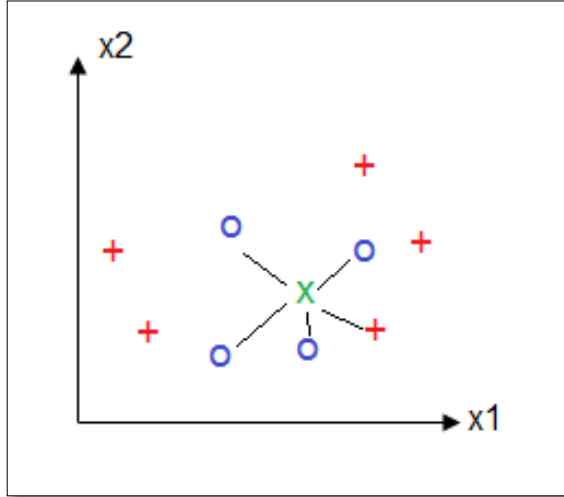
ÇKYSA'da, eğitim örnekleri ağa teker teker sunulur ve başlangıçta öngörülen hata miktarı dahilinde çıkış üretebilecek ağırlık değerlerinin hesaplanması gerçekleştirilir. Başlangıçta tüm değerler rastgeledir. Daha sonra belirtilen kriterlere göre, bağlantı ağırlıkları uygun biçimde güncellenir. İleri yönde hesaplama ile ağ çıkışı bulunup, ortaya çıkan hata değeri dahilinde geriye doğru hesaplama ile ağırlıklar değiştirilir. Hata değeri belli bir seviyenin altına inene kadar öğrenme işlemi devam eder. Öğrenme işleminin tamamlanmasıyla birlikte, ağa test verileri sunulur ve ağın sınıflandırma başarısı belirlenir.

Hata belli bir değer altına düşmüyorsa, ağın öğrenmesi istenen şekilde gerçekleşemez. Hata değerinin istenen değere çekilememesinin sebepleri yanlış örnek seçimi, ağa yönelik parametrelerin (momentum katsayısı, öğrenme katsayısı, aktivasyon

fonksiyonu) yanlış seçilmiş olması, başlangıç değerlerinin uygun aralıkta olmaması, katman veya katmanlardaki eleman sayısının yeterli olmaması gibi parametrelerdir.

1.8.2. k-En Yakın Komşuluk

k-en yakın komşuluk algoritması, örüntü tanıma ve otomatik öğrenmeye ilişkin uygulamalarda, verilerin sınıflandırılmasında kullanılan bir danışmanlı öğrenme metodudur [49]. Bu yöntemde göre, tanıma yapılacak olan örüntü örneğine en yakın olan k komşusu tespit edilir. Daha sonra bu k tane komşunun en fazla hangi sınıfa ait olduğu bulunur ve o sınıfın etiket değeri, tanınacak örüntü örneğine atanır. Bu durumu açıklayan bir örnek, Şekil 13'te verilmiştir. Burada, iki boyutlu uzayda (x_1, x_2) noktalarıyla temsil edilebilen “+” ve “o” sembolleriyle ifadelendirilmiş iki sınıfın dağılımı ve yine bu uzayda bir noktaya karşılık gelen x örüntüsü verilmiştir. x örüntüsü, tanınacak örüntüyü ifade etmektedir. $k=5$ alındığı varsayılınsın. Bu durumda tanınacak olan örüntüye en yakın 5 noktaya bakılır. Bu noktalardan 4 tanesi “+” ile temsil edilen sınıfa, 1 tanesi ise “o” ile temsil edilen sınıfa ait olduğundan x örüntüsüne, “+” ile gösterilen sınıfın etiket değeri verilir.



Şekil 13. k-EYK algoritması örneği

Burada kullanılan k komşuluk parametresi, elde edilecek sınıflandırma doğruluğunu etkileyen birincil parametredir. k parametresinin seçilmesinde de çeşitli metotlar ileri

sürülmüştür. Uygulama kolaylığı nedeniyle genelde deneme-yanılma ya da çapraz doğrulama yöntemleri kullanılmaktadır.

İki nokta arasındaki benzerliği ya da uzaklığı hesaplayacak olan uzaklık hesaplama metodudur ve literatürde çok sayıda uzaklık hesaplama yöntemi vardır (Öklit, Mahalanobis, Manhattan, Chebshev vb.). Literatürde bu yöntemlerin birbirine üstünlüğünü ortaya koyan çok sayıda çalışma da mevcuttur. Ancak genel olarak iki nokta arasındaki uzaklık hesaplanmak istediğinde en yaygın kullanılan yöntem Öklid uzaklığı yöntemidir. Uzayda N boyutlu olduğu varsayılan A ve B gibi iki noktanın uzaklığı, (45) eşitliği ile hesaplanır.

$$d = \sqrt{(A_1 - B_1)^2 + (A_2 - B_2)^2 + \dots + (A_N - B_N)^2} \quad (45)$$

1.8.3. Naive Bayes Sınıflandırıcı

Bayes sınıflandırıcı bir istatistiksel danışmanlı öğrenme metodudur [50]. Bu sınıflandırma teoremi, ismini Bayes teoreminin öncüsü olan Thomas Bayes (1702-1761)'den almıştır. Bayes sınıflandırıcılar, Bayes teoremini kullanırlar. Bayes teoremi, örneklerin hangi sınıfa hangi olasılıkla ait olduklarını öngörmektedir. (46) eşitliği, Bayes teoremini ifade etmektedir.

$$p(c_j|d) = \frac{p(d|c_j)p(c_j)}{p(d)} \quad (46)$$

Burada $p(c_j|d)$, d örneğinin c_j sınıfı içerisinde bulunma olasılığıdır ve hesaplanmak istenen değer budur. $p(d|c_j)$, c_j sınıfı içerisinde d örneğini üretme olasılığıdır; $p(c_j)$, c_j sınıfının meydana gelme olasılığıdır ve bu değer veri tabanında c_j sınıfının ne sıklıkta olduğunu verir. $p(d)$, d örneğinin meydana gelme olasılığıdır ve tüm sınıflar için aynı olduğundan dolayı genellikle ihmal edilir [51]. Sınıflandırıcı kavramına geçmeden önce, Bayes teoreminin çalışmasını ve matematiksel analizini bir örnekle açıklamak gerekirse; 2 tane sınıf olduğu düşünölsün. Bu sınıflar sırasıyla s_1 =erkekler, s_2 =kadınlar olarak ayrılmıştır. Cinsiyeti bilinmeyen bir kişi ve ismi “unisex” yani hem erkek hem de kadın

için kullanılan bir isim olan “Deniz” olsun. Bu durumda bu kişinin kadın ve erkek olma olasılıklarını bulmak için $p(\text{erkekler}|\text{Deniz})$ ve $p(\text{kadınlar}|\text{Deniz})$ olasılıklarına bakılır ve değeri büyük olana göre kadın veya erkek olduğu kararına varılır. Erkekler kümesine ait denklem aşağıda verilmiştir.

$$p(\text{erkekler}|\text{Deniz}) = \frac{p(\text{Deniz}|\text{erkekler})p(\text{erkekler})}{p(\text{Deniz})}$$

$p(\text{Deniz}|\text{erkekler})$, erkek olduğu bilindiğinde Deniz olma olasılığı,

$p(\text{erkekler})$, kümede erkeklerin olma olasılığı,

$p(\text{Deniz})$, Deniz olma olasılığını verir. Benzer şekilde aynı işlem kadınlar kümesi için yapılırsa;

$$p(\text{kadınlar}|\text{Deniz}) = \frac{p(\text{Deniz}|\text{kadınlar})p(\text{kadınlar})}{p(\text{Deniz})}$$

$p(\text{Deniz}|\text{kadınlar})$, kadın olduğu bilindiğinde Deniz olma olasılığı,

$p(\text{kadınlar})$, kümede kadınların olma olasılığı

$p(\text{Deniz})$, Deniz olma olasılığını verir.

Tablo 1'deki gibi bir veri kümesi olsun. Buna göre yukarıdaki ifadelerin değerleri elde edilmek istenirse;

Tablo 1. Bayes için örnek veri seti

İsim	Cinsiyet
Cemal	Erkek
Funda	Kadın
Elif	Kadın
Deniz	Erkek
Nihal	Kadın
Deniz	Kadın
Deniz	Kadın
Caner	Erkek

$$p(\text{erkekler}|\text{Deniz}) = \frac{\frac{1}{3} * \frac{3}{8}}{\frac{3}{8}} = \frac{0.125}{\frac{3}{8}}$$

$$p(\text{kadınlar}|\text{Deniz}) = \frac{\frac{2}{5} * \frac{5}{8}}{\frac{3}{8}} = \frac{0.25}{\frac{3}{8}}$$

Buradan görülmektedir ki, tanınacak olan kişinin kadın olması daha muhtemeldir. Bu örnekte iki sınıfı birbirinden ayırmada tek öznitelikten faydalanıldı. Ancak öznitelik sayısı gerçek problemlerde birden fazla olabilir ve olması gereken durumlar da vardır. Bu durumda ise, bütün durumları birbirinden bağımsız olarak düşünen Naive Bayes sınıflandırıcı kullanılmaktadır. Naive Bayes sınıflandırıcı, niteliklerin birbirlerini etkilemediği ve aynı önem derecesine sahip olduğu düşünülerek ortaya atılan bir sınıflandırıcıdır [52]. n tane öznitelikle temsil edilmiş c_j sınıfının d örneğini üretme olasılığı (47)'deki gibi ifade edilmektedir.

$$p(d|c_j) = p(d_1|c_j) * p(d_2|c_j) * \dots * p(d_n|c_j) \quad (47)$$

Gelen veriler doğrultusunda maksimum olasılığa sahip olan durum, karar durumu olarak seçilir.

Tablo 1'de belirtilen veri setine bir takım öznitelikler eklenerek oluşturulan Tablo 2'deki veri kümesi kullanılarak Naive Bayes sınıflandırıcı örneklenecek olursa;

Tablo 2. Naive Bayes uygulaması için örnek veri kümesi

İsim	Cinsiyet	Yaş	Boy (cm)
Cemal	Erkek	27	185
Funda	Kadın	25	170
Elif	Kadın	27	168
Deniz	Erkek	36	170
Nihal	Kadın	25	168
Deniz	Kadın	32	160
Deniz	Kadın	29	165
Caner	Erkek	22	192

Bilinen, Deniz kişinin boyunun 165 cm ve üzeri yaşının 30'dan küçük olduğudur. Bu durumda cinsiyetini belirlemek için;

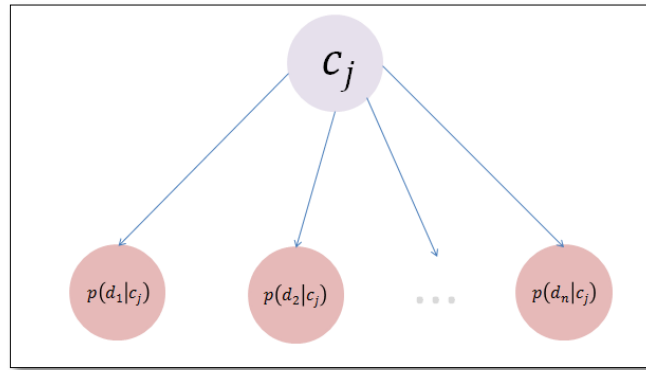
$$p(\text{Deniz} \mid \text{kadınlar}) = p(165 \text{ ve üstü} \mid \text{kadınlar}) * p(30' \text{ dan küçük} \mid \text{kadınlar})$$

$$p(\text{Deniz} \mid \text{erkekler}) = p(165 \text{ ve üstü} \mid \text{erkekler}) * p(30' \text{ dan küçük} \mid \text{erkekler})$$

$$p(\text{Deniz} \mid \text{kadınlar}) = \frac{4}{5} * \frac{4}{5}$$

$$p(\text{Deniz} \mid \text{erkekler}) = \frac{3}{3} * \frac{2}{5}$$

Buradan görüleceği gibi, bu özelliklere sahip Deniz kişinin cinsiyetinin kadın olduğu bilgisi kesin olarak elde edilmiştir. Naive Bayes sınıflandırıcılar, Şekil 14'te gösterilen graf yapısı ile temsil edilirler.



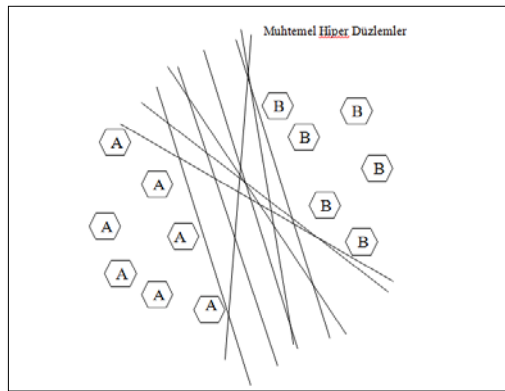
Şekil 14. Naive Bayes sınıflandırıcının graf yapısı

Naive Bayes sınıflandırıcının birtakım avantaj ve dezavantajları bulunmaktadır. Naive Bayes sınıflandırıcı sayesinde tüm olasılıklar veri tabanının tek bir taranması ile elde edilir ve tüm bu olasılıklar küçük bir tabloda saklanabilir. Bu sebeple hızlı ve uzay bakımından etkin bir yöntemdir. Bir diğer avantajı ise, Naive Bayes sınıflandırıcı, aranılan özellikle ilişkisiz özniteliklere karşı duyarlı değildir ve bu da olasılık değerini etkilemeyeceğinden, sınıflandırma işlemi başarısız olmayacaktır [53]. Örneğin Tablo 2'de, "göz rengi" şeklinde bir öznitelik olsaydı, ki bu öznitelik bir kişinin cinsiyetiyle tamamen ilişkisizdir, bu durumda da sınıflandırıcı etkin bir şekilde çalışabilmektedir.

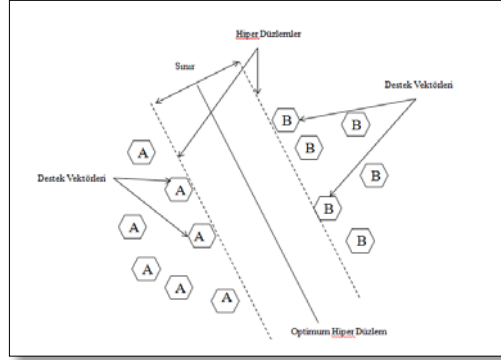
Dezavantajlarına değinilecek olursa, bundan önceki örnekte iki sınıf ve iki örnekle sınıflandırma yapılabilmekteydi. Ancak çok sınıflı ve çok öznitelikli problemlerde, Naive Bayes sınıflandırıcının öznitelikleri birbirinden bağımsız ilişkilendirmesi mantığı yanlış sınıflandırmalara sebebiyet verebilmektedir. Çünkü gerçek hayatta değişkenler birbirine bağımlıdır. Buna, öznitelikler arasında ilişki olduğu düşünülerek çözüm getirilebilir. Bunu gerçekleyen yapıya ise “Bayes Ağları” denir [54].

1.8.4. Destek Vektör Makineleri

Destek vektör makineleri, danışmanlı öğrenme yaklaşımıyla çalışan bir sınıflandırma tekniğidir. Hem doğrusal ayrılabilen hem de doğrusal ayrılamayan sınıflandırma problemlerinde kullanılır. Ancak DVM'nin çalışma prensibi temel olarak iki sınıfa ait verileri birbirinden en uygun şekilde ayıran karar sınırlarının yani hiper-düzlemlerin belirlenmesine dayalıdır [55], [56]. Doğrusal olarak ayrılabilen iki sınıflı bir verinin sınıflandırılması, DVM için en temel sınıflandırma problemidir. Sınıf etiketleri bilinen iki sınıfa ait örneklerin, eğitim ve test verileri olacak şekilde ayrılıp; eğitim verisi sonucu elde edilen bir karar fonksiyonu sayesinde test verilerinin sınıflandırılması işlemidir. Bahsi geçen karar fonksiyonu kullanılarak, verileri en iyi şekilde ayıran hiper-düzlem bulunur [57]. Şekil 15'e bakıldığında iki sınıfı birbirinden ayıran birden çok hiper-düzlem olabilir. Ancak destek vektör makinelerindeki temel amaç, kendisine en yakın noktalar arasındaki uzaklığı maksimuma çıkaran hiper-düzlemin bulunmasıdır. Şekil 16'da gösterilen, sınırı maksimuma çıkarıp en uygun ayrımı yapan hiper düzleme optimum hiper-düzlem ve sınır genişliğini sınırlandıran noktalara ise destek vektörleri denir.



Şekil 15. İki sınıflı problemler için muhtemel hiper-düzlemler



Şekil 16. Optimum hiper-düzlem ve destek vektörleri

1.8.4.1. Doğrusal Ayrılabilen Veriler İçin DVM

DVM'nin en temel ve basit modeli, doğrusal ayrılabilen sınıflar için geçerli modelidir. Bu model aynı zamanda doğrusal olmayan ayrıştırma için de temel oluşturur [58].

Doğrusal olarak ayrılabilen, $y=\{-1,1\}$ etiketlerine sahip iki sınıflı ($\{w_1, w_2\}$) bir veri kümesi düşünölsün. k sayıda örnekten oluşan bir eğitim örneđi seti ($\{x_i, i=1, \dots, k\}$) olsun. Doğrusal ayırma fonksiyonu (48) eşitliđi ifade edilir.

$$g(x) = w^T x + w_0 \quad (48)$$

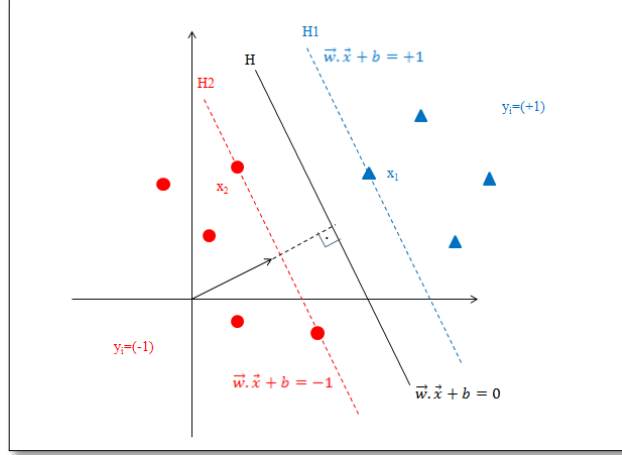
Bu eşitlik, karar fonksiyonunu bulmak için düzenlendiđinde (49)'daki gibi bir denklem sistemi elde edilir.

$$w^T x + w_0 = \pm 1 \quad (49)$$

Bu ifade tek bir eşitsizlik olarak ifade edildiđinde, tüm eğitim örnekleri için (50) eşitsizliđi sağlanmış olmalıdır.

$$y_i (w^T x_i + w_0) > 0 \quad (50)$$

Verilen eğitim kümesini doğrusal olarak ayırabilecek olan en büyük sınırlı hiper-düzlemin geometrik sınırı, Şekil 17'de gösterilen düzleme göre (51), (52), (53), (54), (55), (56) ve (57) ifadelerinden hesaplanır. Bulunan bu en büyük sınırı "marjin" denir.



Şekil 17. Doğrusal ayrıştırılabilir veriler için hiper-düzlem ve destek vektörleri (*sert marjin*)

$$w \cdot x + b \geq +1, y_i = +1 \quad (51)$$

$$w \cdot x + b \leq -1, y_i = -1 \quad (52)$$

$$y_i (w \cdot x + b) \geq +1 \quad (53)$$

$$w \cdot x_1 + b = y_i = +1 \quad (54)$$

$$w \cdot x_2 + b = y_i = -1 \quad (55)$$

$$w(x_1 - x_2) = 2 \quad (56)$$

$$\frac{w}{\|w\|} (x_1 - x_2) = \frac{2}{\|w\|} \quad (57)$$

Optimum hiper-düzlemin sınırının maksimuma çıkarılabilmesi için, $\|w\|$ ifadesinin minimum hale getirilmesi gerekir. Bunun içinse, (58)'deki sınırlı optimizasyon probleminin belirtilen sınırlamalar dahilinde çözülmesi gereklidir.

$$\min \left[\frac{1}{2} \|w\|^2 \right], \quad y_i (w \cdot x_i + b) - 1 \geq 0 \quad \text{ve} \quad y_i \in \{1, -1\} \quad (58)$$

Bu optimizasyon probleminin çözümü ise Lagrange denklemleri metodu ile mümkündür. Bu işlem sonrasında (59) eşitliğine geçilir.

$$L(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^k a_i y_i (w \cdot x_i + b) + \sum_{i=1}^k a_i \quad (59)$$

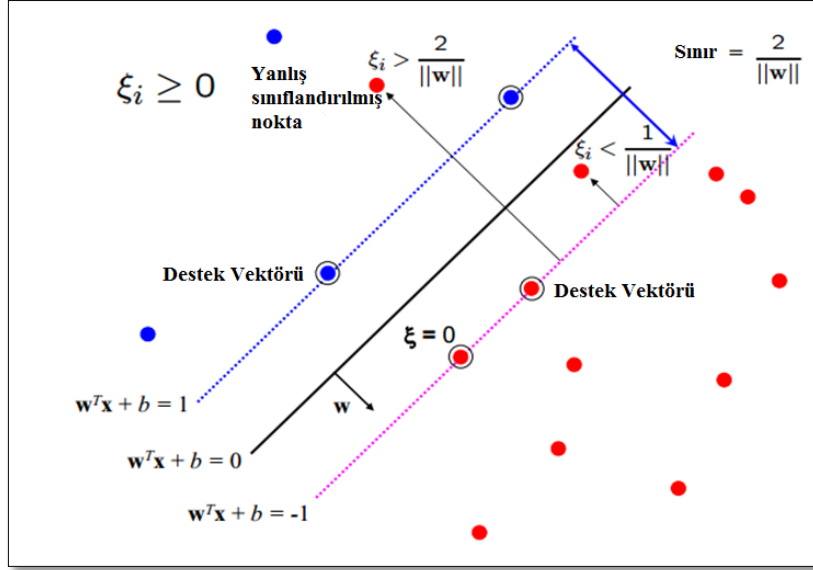
Sonuç olarak doğrusal olarak ayrılabilen iki sınıflı problemler için (60) eşitliği elde edilir.

$$f(x) = \text{sign} \left(\sum_{i=1}^k \lambda_i y_i (x \cdot x_i) + b \right) \quad (60)$$

Doğrusal ayrıştırma için bir diğer yaklaşım “yumuşak marjin” denilen yaklaşımdır. Bu durumda, eğitim verilerinin bir kısmının, belirlenen optimum hiper-düzlemin karşı tarafında kalması problemi bir yapay değişken ξ_i tanımlanması ile çözülebilir. Sınırın maksimum hale getirilmesi ve sınıflandırma hatalarının minimize edilmesi arasındaki ilişki, bir düzenleme parametresi (C , ($0 < C < \infty$)) ile kontrol edilebilir [59]. Düzenleme parametresi ve yapay değişken ile optimizasyon ifadesi yeniden düzenlenir ve (61) şeklini alır. Buna bağlı kısıtlamalar ise (62)'de ifade edilmiştir.

$$\min \left[\frac{\|w\|^2}{2} + C \cdot \sum_{i=1}^r \xi_i \right] \quad (61)$$

$$y_i (w \cdot \varphi(x_i) + b) - 1 \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad \text{ve} \quad i = 1, \dots, N \quad (62)$$



Şekil 18. Doğrusal ayrılabilen veriler için yumuşak marjin

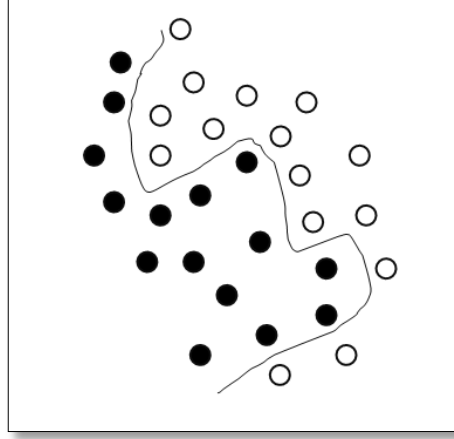
Şekil 18'de, $0 < \xi \leq \frac{1}{\|w\|}$ olduğu durumda nokta, marjin (sınır) aralığındadır ve

hiper-düzlemin doğru tarafındadır. $\xi \geq \frac{1}{\|w\|}$ için ise nokta yanlış sınıflandırılmıştır.

Ancak yumuşak marjin de her problemin çözümü için yeterli olmamaktadır. Sonuç itibariyle belli bir oranda nokta hiper-düzlemin karşı tarafında kalmalıdır. Bu nedenle doğrusal olmayan ayrıştırma için geliştirilen, çekirdek tabanlı DVM uygulamaları mevcuttur [57].

1.8.4.2. Doğrusal Ayrılamayan Veriler için DVM

Gerçek zamanlı uygulamaların çoğu için, doğrusal ayrıştırmanın mümkün olmadığı söylenebilir. Şekil 19'da da buna benzer bir durum söz konusudur. Burdan görüleceği üzere, girdi uzayında doğrusal olarak ayrılamayan veriler, girdi uzayından daha yüksek boyutlu olan bir başka uzayda tanımlanır. Bu uzaya öznitelik uzayı da denir. Böylece veriler doğrusal olarak ayrılabilir ve hiper-düzlem belirlenebilecek duruma getirilir. Şekil 19'da doğrusal ayrıştırılmayan bir problem verilmiştir.



Şekil 19. Doğrusal ayrıştırılmayan bir sınıflandırma problemi

DVM’de, çekirdek fonksiyonu denilen fonksiyonlar sayesinde doğrusal olmayan dönüşümler yapılarak, verilerin bulunduğu uzayın boyutundan yüksek bir boyutta doğrusal olarak ayrılması imkânı sağlanmaktadır. Bununla ilgili matematiksel ifade (63)’te verilmiştir.

$$f(x) = \text{sign} \left(\sum_i a_i y_i \varphi(x) \cdot \varphi(x_i) + b \right) \quad (63)$$

Doğrusal uygulamalarda giriş verileri iç çarpım ile sisteme verilirken, karmaşık uygulamalarda, giriş verileri bir çekirdek fonksiyonundan geçirilerek uzaya düşürülür [60]. Bu fonksiyon ifadedeki φ 'ye karşılık gelir.

DVM ile gerçekleştirilecek bir sınıflandırma problemi için kullanılacak çekirdek fonksiyonu ve buna dair optimum parametrelerin belirlenmesi önemli adımlardan biridir. Literatürde en yaygın kullanılan çekirdek fonksiyonları olan polinom, radyal tabanlı fonksiyon, Pearson VII fonksiyonu ve normalleştirilmiş polinom çekirdekleri, Tablo 3’te formül ve parametreleriyle birlikte verilmiştir.

Tablo 3. Destek vektör makinelerinde kullanılan temel çekirdek fonksiyonları ve parametreleri

Çekirdek Fonksiyonu	Matematiksel Formülü	Parametreler
Polinom Çekirdek Fonksiyonu	$K(x, y) = ((x \cdot y + 1)^d$	Polinom derecesi (d)
Normaleştirilmiş Polinom Çekirdek Fonksiyonu	$K(x, y) = \frac{((x \cdot y + 1)^d}{\sqrt{((x \cdot x + 1)^d((y \cdot y + 1)^d}}$	Polinom derecesi (d)
Pearson VII (PUK) Çekirdek Fonksiyonu	$\frac{1}{\left[1 + \left(\frac{2\sqrt{ x - y ^2 \sqrt{2^{1/\omega} - 1}}}{\sigma} \right)^2 \right]^\omega}$	Pearson genişliği parametreleri (σ, ω)
Radyal Tabanlı Çekirdek Fonksiyonu	$K(x, y) = e^{-\gamma x - x_i ^2}$	Çekirdek boyutu (γ)

Çekirdek fonksiyonları karşılaştırıldığında, genellikle polinom ve radyal tabanlı çekirdeklerin daha anlaşılır ve sade olduğu ve uygulamalarda daha çok tercih edildiği görülmektedir. [61], [62] ve [63] çalışmaları bunlardan bazılarıdır.

Bunların yanı sıra, tüm destek vektör makineleri için düzenleme parametresi C'nin kullanıcı tarafından belirlenmesi gerekir. Bu parametre için olması gerekenden çok küçük veya çok büyük değerler seçilmesi durumunda optimum hiper-düzlem doğru belirlenemeyeceğinden sınıflandırma doğruluğunda ciddi düşüş beklenir. C= ∞ olması durumunda, DVM sadece doğrusal ayrıştırılabilen veri setleri için uygulanabilir hale gelir [59].

Tüm bunlardan yola çıkılarak, DVM sınıflandırıcısının performansını değerlendirmede parametre belirlemenin ne kadar önemli olduğu görülmektedir. Genellikle deneme-yanılma stratejisi tercih edilse de, çapraz doğrulama yaklaşımı etkin sonuçlar elde edilmesine imkân sağlamaktadır.

1.8.4.3. DVM ile Çoklu Sınıflandırma

Çoklu sınıflandırma, ikiden fazla sınıfın ayrılması durumunda gerekmektedir. Bu tür problemlerde ise DVM doğrudan sınıflandırma için kullanılamamaktadır. Çünkü bu sınıflandırıcı yaklaşımı, iki sınıflı problemler üzerine kurulmuştur. Bu sebeple DVM'yi çok

sınıflı problemlerde de kullanabilmek amacıyla birkaç yaklaşım ortaya atılmıştır. Bunlardan en sık kullanılanları; bire-karşı-hepsi (one vs. all), bire-karşı-bir (one vs. one) ve böl ve fethet (divide and conquer) yaklaşımlarıdır.

Bire-karşı-hepsi yöntemi, kavramsal olarak en temel çoklu DVM yöntemidir. Burada k tane sınıf olduğu düşünülür. Sınıf 1 pozitifken, diğer tüm sınıflar negatif; Sınıf 2 pozitifken, diğer tüm sınıflar negatif ve en son Sınıf k pozitifken, diğer tüm sınıflar negatif olarak düşünülür. Buna bağlı olarak her bir ikili için kullanılan k tane sınıflandırıcı ile eğitim kümesi eğitilir. Bir test örneği, bu ikili sınıflandırma sonuçlarına göre bir sınıfa dahil olur. Sınıflandırıcı sonucunun pozitif olması test örneğinin i . sınıfa dahil olduğunu, sonucun negatif olması ise test örneğinin bunun dışında kalan sınıflara dahil olduğunu ifade eder. Ancak sonucun pozitif veya negatife bağlı olma durumu, ilerleyen aşamalarda bazı eşit durumların çıkmasına sebep olmaktadır. Buna sebep olan durumlar ifade edilecek olursa; iki ya da daha fazla sınıflandırıcının çıkışının pozitif olması durumu veya her sınıflandırıcının çıkışının negatif olması durumu örnek verilebilir. Bunu gidermek için algoritmaya bir takım yaklaşımlar ilave etmek gerekmektedir [64].

Bire-karşı-bir yönteminde, k tane kümenin ikili kombinasyonları bulunur. Bunlar sayıca $M(M-1)/2$ tanedirler. Bunların her birisi eğitilir ve bir test örneği için her bir sınıflandırıcı, test örneğinin hangi sınıfa dâhil olması gerektiğine dair bir bilgi verir. Bu durum, oylama olarak da düşünülebilir. En fazla oyu alan sınıf, test örneğinin dâhil olacağı sınıftır. Ancak bu yöntemde de, bire-karşı-hepsi yönteminde olduğu gibi eşitlik durumları olabilir ve bire-karşı-hepsi yöntemine benzer şekilde giderilebilir [64].

Böl ve fethet yönteminde ise, karar ağacı tabanlı bir yaklaşım öne sürülmüştür. Bu yöntemin sınıflandırma doğruluğunu hesaplamak için, en büyük bilgi kazancını sağlayan özniteliklerden yola çıkılır. Bunun için de, eğitim seti küçük alt kümelere bölünür ve bu bölme işlemi her bir alt kümede bir tek sınıf kalana dek devam eder. Sınıf sayısı arttıkça, oluşacak alt küme olasılıkları da artacaktır. Dolayısıyla bu yaklaşımın değerlendirilmesi, bir önceki yöntemlere göre daha fazla yük getirecektir. Ancak diğer yöntemlerden avantajlı olarak, bu yöntemde eşitlik çıkması durumu yoktur.

Çalışmada uygulama kolaylığı açısından bire-karşı-bir yaklaşımı kullanılarak sınıflandırma performansı hesaplanmıştır [65].

2. YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME

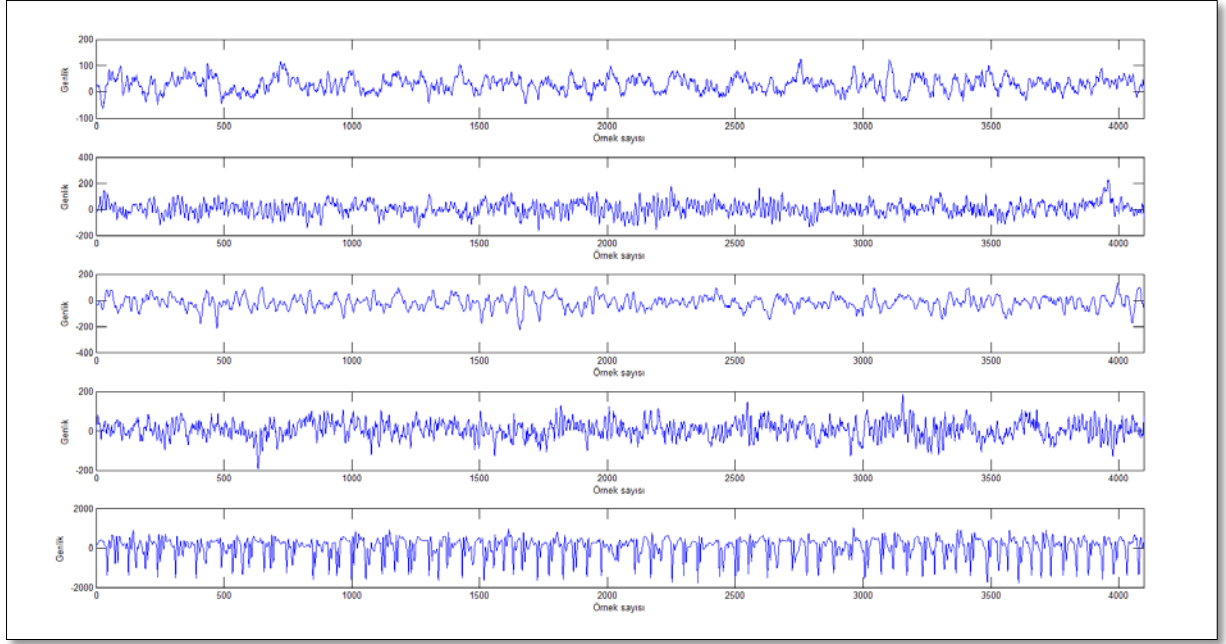
2.1. Çalışmanın Yapısı ve Adımları

Bu tez çalışmasında epilepsi hastalarının epileptik nöbet dönemlerinin ayırt edilmesi amacıyla dayanan sistemin temeli oluşturulmaktadır. Mevcut veri setiyle birlikte kişinin sağlıklı, nöbet olmayan ve nöbet dönemlerinin ayırt edilmesi işlemi gerçekleştirilmektedir.

2.2. Kullanılan Veri Seti

Bu tez çalışmasında Almanya Bonn Üniversitesi, Epileptoloji Bölümü'nde Andrzejak vd. tarafından epilepsi teşhisi ve epileptik nöbet belirleme araştırmalarında kullanılmak üzere kaydedilen ve ortak kullanıma açık EEG veri seti kullanılmıştır [10], [66]. Bu veri seti 128-kanallı 12 bit EEG kayıt sistemi ile, 173.61 örnekleme frekansıyla örneklenecek kaydedilmiştir, bu 128 kanalın ortalama değeri alınmıştır. Toplam 500 denemeden oluşan bu sistem, 5 grup bireyden elde edilmiştir (A, B, C, D ve E). Her bir deneme 23.6 sn uzunluğundadır. Dolayısıyla her bir deneme $173,61 \times 23,6 = 4097$ örnekle temsil edilir. Tüm denemeler, göz ve kas hareketlerinden dolayı oluşan artifaktlardan arındırılmıştır.

A ve B alt grupları, 5 gönüllü sağlıklı bireyden uluslararası 10-20 elektrot sistemi yerleştirilerek, sırasıyla göz açık ve kapalı durumda kaydedilmiştir. C ve D alt grupları 5 epilepsi hastası bireyden nöbet olmayan dönemde ölçülen ve kafatası içinden alınan kayıtlardır. Bu kayıtlar için, elektrotlar C ve D alt grupları için sırasıyla epileptik odağa ve karşıt yarım kürelerin beyin çıkıntısına yerleştirilmiştir. Son olarak E alt grubu, aynı 5 hastada bireyden alınan nöbet kayıtlarıdır [10]. Bu 5 alt gruba ait birer örnek deneme; A, B, C, D ve E sırasıyla Şekil 20'de verilmiştir.



Şekil 20. A,B,C,D ve E gruplarına ait örnek EEG verileri

Her bir alt grubun belirli kombinasyonları ile, üzerinde çalışılacak olan veri grupları elde edilmiştir. Tablo 4'te, oluşturulan veri kombinasyonları ve açıklamaları verilmiştir.

Tablo 4. Üzerinde çalışılacak olan veri grupları

Alt Gruplar	Sınıflar			Açıklama
Küme 1	Sınıf 1 A+B	Sınıf 2 C+D	Sınıf 3 E	Veri seti üç sınıfa ayrılarak, sağlıklı, nöbet öncesi ve hastalıklı dönemin sınıflandırılması için hazırlanmıştır.
Küme 2	Sınıf 1 A	Sınıf 2 D	Sınıf 3 E	Sağlıklı insanda göz açıkkenki ayırt ediciliği tespit etmek açısından hazırlanmıştır.
Küme 3	Sınıf 1 B	Sınıf 2 D	Sınıf 3 E	Sağlıklı insanda göz kapalıykenki ayırt ediciliği tespit etmek açısından hazırlanmıştır.
Küme 4	Sınıf 1 A+B	Sınıf 2 E		Yalnızca sağlıklı ve hastalıklı dönemin birbirinden ayrılması için hazırlanmıştır.
Küme 5	Sınıf 1 A+B+C+D	Sınıf 2 E		Hastalıklı dönemin diğer tüm durumlardan ayrılması amaçlanmıştır.
Küme 6	Sınıf 1 A+B	Sınıf 2 C+D+E		Hasta ve sağlıklı bireylerin ayrılması için hazırlanmıştır.
Küme 7	Sınıf 1 A+B	Sınıf 2 C+D		Nöbet geçirmeme durumunda, hasta ve sağlıklı bireyin EEG'lerinin birbirinden farklılığını ortaya koymak adına oluşturulmuştur.

2.3. Özniteliklerin Elde Edilmesi

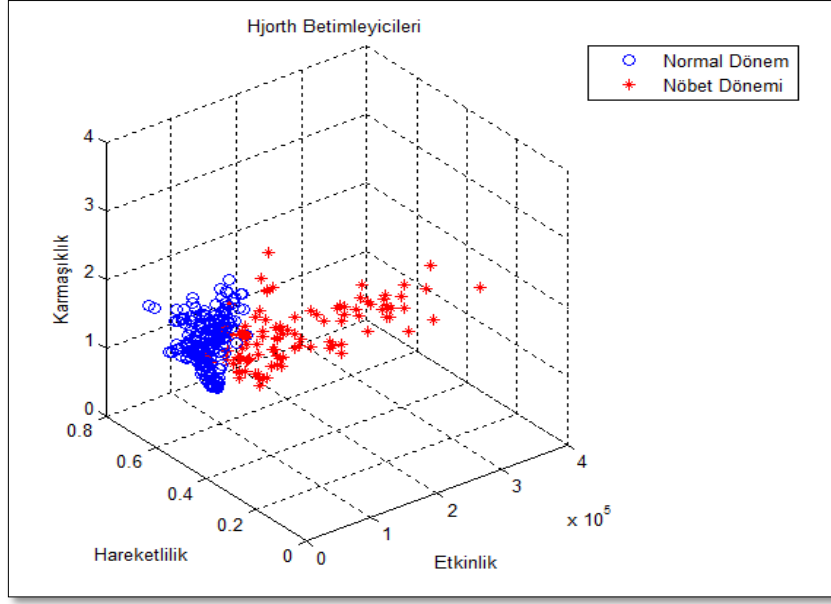
Öznitelik çıkarımında zaman uzayı yöntemlerinden faydalanılmıştır. Oluşturulan tüm alt gruplar önce çapraz doğrulamanın 10-kat doğrulama yöntemi ile %60 eğitim ve %40 test verisi olacak şekilde rastgele olarak ayrılıp, daha sonra eğitim ve test verileri için öznitelikler çıkarılmış ve test işlemleri ile her bir grup için ortalama sınıflandırma doğruluğu bulunmuştur.

2.3.1. Hjorth Betimleyicilerinin Uygulanması

EEG işaretinin zaman uzayında özneliğinin çıkarılmasında ilk olarak Hjorth betimleyicilerinden faydalanılmıştır. Literatürdeki başarılarından yola çıkılarak, sağlıklı ve epileptik sinyallerin ayrıştırılması için uygulanmıştır.

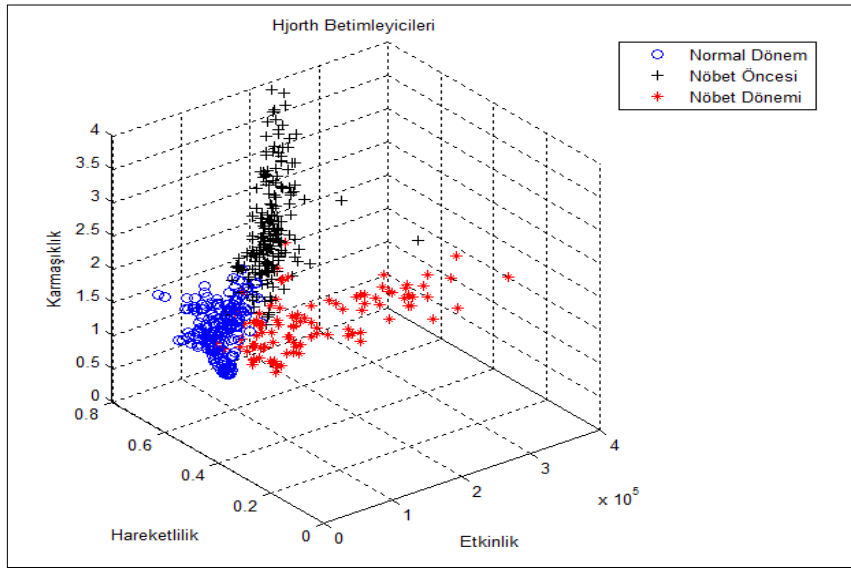
Hjorth betimleyici yöntemi, zaman uzayında doğrusal olmayan ayrıştırma parametrelerinden birisidir. Birçok çalışmada Hjorth betimleyicilerinden nöbet teşhisi, alt bantlara ayrıştırma gibi amaçlarla faydalanılmıştır. Aynı zamanda EEG işaretlerinin zaman uzayında temsil edilmesinde Hjorth betimleyicilerinin tutarlılığını ve güvenilirliğini araştıran, BBA sistemlerine yönelik çalışmalar da mevcuttur. Bunlardan biri, [6] çalışmasında EEG sinyalleri üzerinden epileptik nöbetin araştırılmasına yöneliktir. Cecchin vd. Hjorth betimleyicilerinden iki tanesini kullanarak EEG sinyalinin karesel ortalamasını ve baskın frekansını hesaplamış, %96 civarında bir başarı elde etmişlerdir. Gudmundsson vd. EEG özelliklerini zaman uzayında temsil etmede, zaman uzayı niceliklerinden biri olarak Hjorth betimleyicilerini de ele almış ve güvenilirliğini analiz etmişlerdir [36]. Benzer şekilde, [67] çalışmasında Vidaurre vd. EEG tabanlı BBA sistemleri için, Hjorth betimleyicilerini zaman uzayında bir öznitelik olarak öngörmüşlerdir.

Bu çalışmada kullanılan örnek bir veri kümesi için sağlıklı dönemin nöbet döneminden ayrılmasında Hjorth betimleyicileri kullanılması ve bu parametrelerin uzayda dağılımları Şekil 21’de gösterilmiştir.



Şekil 21. Normal ve epileptik nöbet dönemlerinin Hjorth betimleyicilerine göre uzayda dağılımları

Bu yöntem Şekil 21'e göre, iki sınıfın birbirinden ayrılması için yeterli olabilir. Ancak Şekil 22'e bakılacak olursa, burada üç dönemin birbirinden sağlıklı ayrıştırılması için Hjorth betimleyicileri tek başına yeterli olmamaktadır. Bu da ek yöntem ya da yöntemlerin uygulanmasını gerektirmektedir.



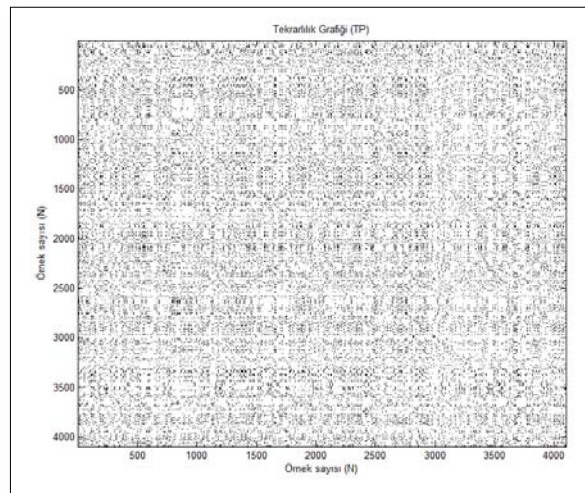
Şekil 22. Normal, nöbet öncesi ve nöbet dönemlerinin Hjorth betimleyicilerine göre uzayda dağılımları

2.3.2. Tekrarlılık Ölçme Analizinin Uygulanması

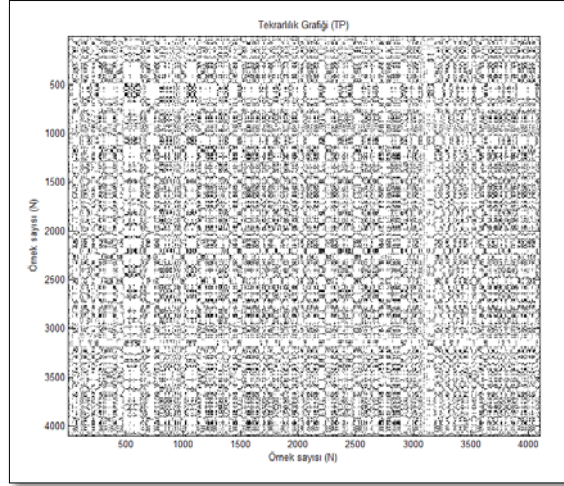
Çalışmada kullanılan bir diğer zaman uzayı yöntemi ise TÖA'dır. Bu yöntem, epilepsi sinyali gibi periyodik ve durağan olmayan sinyallere ilişkin tekrarlılık oranlarını ortaya koyan iki boyutlu bir grafikdir [68].

Tekrarlılık grafiğinin yorumlanmasında tekrarlılık ölçme analizi (TÖA) parametrelerinden faydalanılır. TÖA parametreleri, bir zaman uzayı sinyalindeki durağan olmama ve gizli periyodiklik durumlarını analiz eder. Acharya vd. epileptik nöbetin tespiti için EEG sinyallerinin sınıflandırılmasında TÖA parametrelerinden faydalanmıştır [38], [39]. Yine başka bir çalışmada [69], tek kanaldan kaydedilen epileptik nöbet işaretleri kullanılarak epileptik nöbetin tahmin edilmesinde TÖA parametreleri kullanılmıştır. [70] çalışmasında, sıçanlardan alınan EEG kayıtları ile, epileptik nöbetin otomatik tanınması yapılmıştır. Bir diğer çalışmada, [70] çalışmasına benzer şekilde gibi yine sıçanlardan alınan EEG kayıtlarıyla epileptik nöbet öncesi dönemin dinamik karakteristiklerinin araştırması yapılmıştır [71]. Çalışmalar sadece epilepsi hastalığının teşhisini araştırmakla kalmamış, [72] çalışmasında meditasyon boyunca insanın kalp atışlarının tekrarlılığı TÖA ile analiz edilmiştir. [73] çalışmasında ise, kardiyak aritmaların (ritim bozukluğu) tespiti TÖA kullanılarak yapılmıştır.

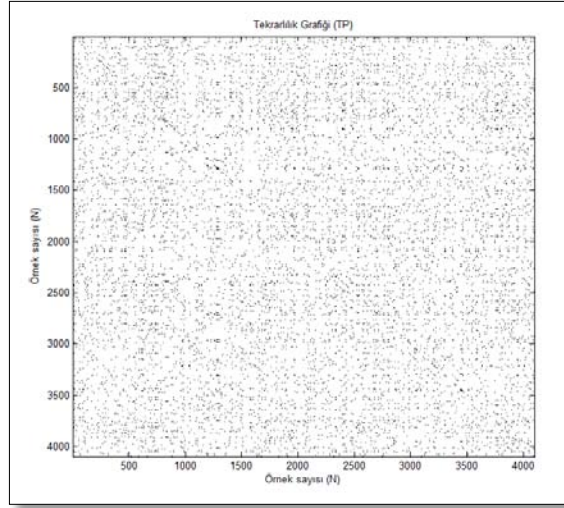
Sırasıyla Şekil 23, 24 ve 25 sağlıklı, nöbet öncesi ve nöbet dönemlerine ilişkin tekrarlılık ölçme grafiklerini vermektedir.



Şekil 23. Sağlıklı döneme ait tekrarlılık grafiği



Şekil 24. Nöbet öncesi döneme ait tekrarlılık grafiği

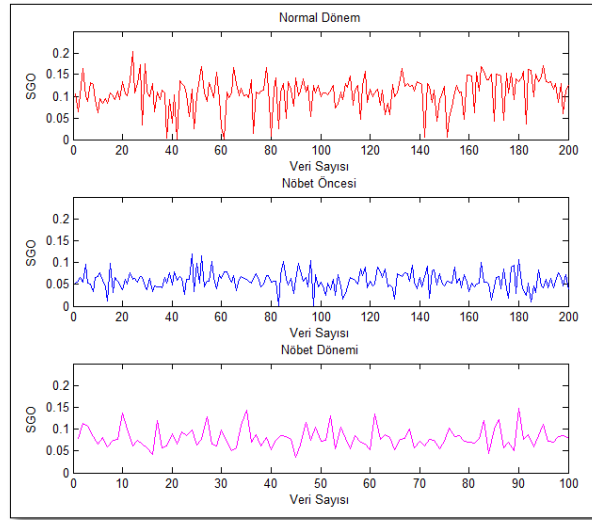


Şekil 25. Nöbet dönemine ait tekrarlılık grafiği

2.3.3. Sıfır Geçiş Oranının Elde Edilmesi

Sıfır geçiş oranı, sinyalin pozitiften negatife ya da negatiften pozitifte geçiş oranı olarak da bilinmektedir [40]. Bu öznelikle genellikle konuşma ve müzik sinyalleri üzerinde kullanılır. Ancak bu uygulamanın EEG veya diğer biyomedikal ya da biyoelektrik sinyallerle olan çalışmaları da mevcuttur. Bunlardan birisi EEG işaretleri kullanarak Alzheimer hastalığının teşhisinin olduğu [41] çalışmasıdır. Bundan başka [74] çalışmasında, Ali Shahidi Zandi vd. kafa derisinden alınan EEG sinyallerinin pozitif sıfır geçiş oranlarını

bulmayı amaçlamışlar ve böylece beyin hareketlerini tanımlayıp epileptik nöbetin teşhisini gerçekleştirmişlerdir. Benzer olarak, [75] çalışmasında Condarsen vd., yine epileptik nöbetin belirlenmesinde sıfır geçiş oranından faydalanmış ancak bu defa elektromiyogram (EMG) işaretlerini kullanmışlardır. Şekil 26'da normal, nöbet öncesi ve epileptik nöbet dönemlerinin SGO dağılımları verilmiştir. Burada normal, nöbet öncesi ve epileptik nöbet dönemleri için ortalama SGO değerleri yaklaşık olarak sırasıyla 0.1, 0.05 ve 0.08 olarak bulunmuştur.



Şekil 26. Normal, nöbet öncesi ve epileptik dönemlere ait SGO dağılımları

2.3.4. Hurst Üstel Katsayısının Elde Edilmesi

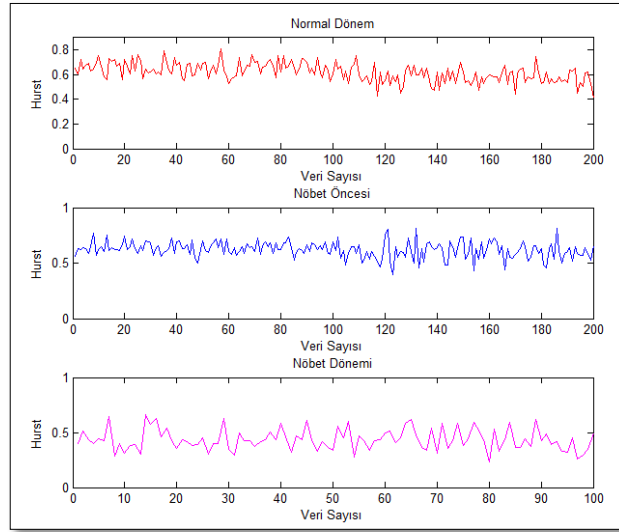
Hurst üstel katsayısı, bir zaman uzayı işaretin kendine benzerlik oranıdır. Harold Edwin Hurst tarafından ilk olarak Nil nehrinin su seviyesinin düzenliliğini araştırmada kullanılan bu metot, günümüzde tıbbi ve finansal uygulamalarda da kullanılmaktadır.

Hurst üstel katsayısını hesaplamak için bir çok yöntem önerilmiştir; ancak klasik olan yöntem, Hurst tarafından geliştirilen “yeniden ölçeklendirilmiş aralık” metodudur ve (64) ifadesiyle hesaplanmaktadır.

$$(R/S)_t = c \cdot t^H \quad (64)$$

Bu formüle göre, Hurst üstel katsayısının elde edilebilmesi için c , sabit katsayısı; R , kümülatif toplamdan sapma miktarı; S , kümülatif toplamın standart sapma değeri ve t , zaman veya gözlem parametresidir. Hurst üstel katsayısına ulaşabilmek için her iki tarafın logaritması alınmış ve logaritmik bir ifade elde edilmiştir.

Literatürde Hurst üstel katsayısının sinyal işleme ve bilhassa EEG sinyallerinin analizinde kullanıldığı birçok çalışma mevcuttur. [42] çalışmasında elektrokortikogram (ECoG) sinyalleri üzerinden epileptik nöbetin algılanmasında Hurst üstel katsayısı ve türlerinden faydalanılmıştır. Bir başka çalışmada [76], insanın uyku sinyallerinin çok parçalı analizinde Hölder ve Hurst üstel katsayısı özneliklerine başvurulmuş ve uyku durumlarının dinamik değerlendirilmesi yapılmıştır. Şekil 27'de mevcut üç döneme ait Hurst üstel katsayısı dağılımları verilmiş ve normal, nöbet öncesi ve nöbet dönemleri için sırasıyla 0.60, 0.62 ve 0.43 değerleri elde edilmiştir. Buradan, normal ve nöbet öncesi dağılımın birbirine yakın sonuçlar verdiği görülmektedir.



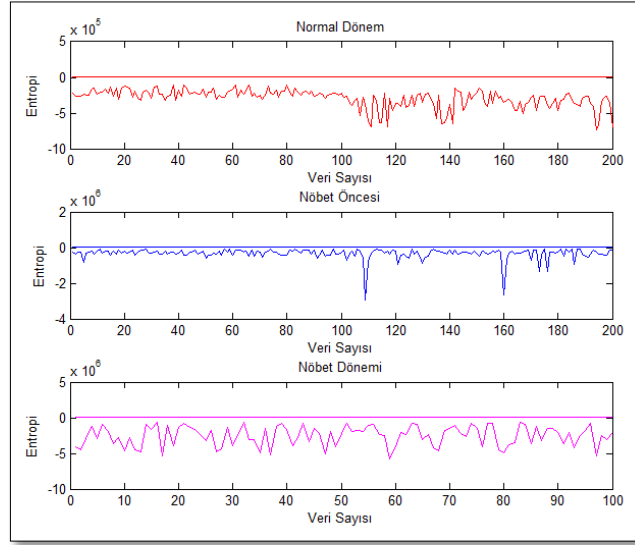
Şekil 27. Normal, nöbet öncesi ve epileptik dönemlere ait Hurst üsteli dağılımları

2.3.5. Entropi Değerinin Elde Edilmesi

Entropi, 1940'lı yılların sonunda Shannon tarafından geliştirilen, bilgi teorisinin temel yapı taşlarından birisidir. Verinin dağılımının bir ölçütüdür. EEG açısından

incelendiğinde, sinyalin düzensizliğinin oranıdır. Eğer sinyal düzensizliği fazla ve tahmin edilebilirliği az ise, yüksek bir değer; düzenli bir sinyal ise düşük bir değer elde edilir.

Bu çalışmada, bahsi geçen üç dönemin birbirinden ayrılmasında entropi ölçütlerinin temeli olan Shannon entropi kuramından faydalanılmıştır. Her bir döneme ait örneklerin Shannon entropi değerleri bulunmuş ve Şekil 28'deki gibi bir dağılım elde edilmiştir.

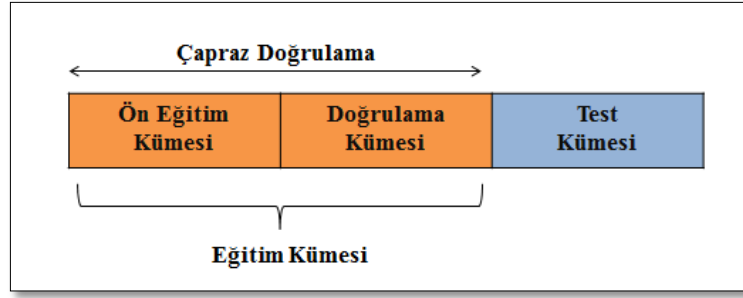


Şekil 28. Normal, nöbet öncesi ve nöbet dönemlerine ilişkin entropi dağılımları

2.4. Çapraz Doğrulama

Çapraz doğrulama (cross validation), öğrenme veya sınıflandırma algoritmalarını karşılaştırma ve değerlendirmeye yarayan istatistiksel bir metottür. Örüntü tanıma problemlerinde parametre seçimi ve performans hesaplanması amaçlarıyla kullanılırlar.

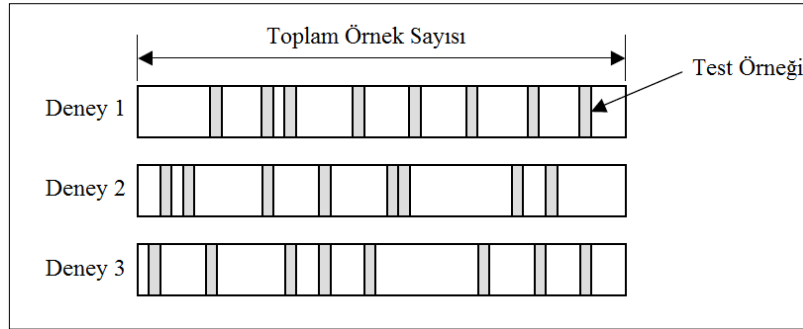
Başlangıçta eğitim ve test verisi olarak ayrılan veri kümesinde, çapraz doğrulama yapabilmek için eğitim verisi de iki gruba ayrılır: Ön eğitim (pre-training) ve doğrulama (validation). Doğrulama verileri eğitim esnasında hesaba katılmaz, eğitim sonuçları doğrulama verileriyle test edilir. Bu sayede öğrenme veya sınıflandırma algoritmaları için en iyi performansı sağlayan parametreler belirlenir ve test işlemi bu değerler dâhilinde yapılarak en yüksek öğrenme ve sınıflandırma sonucu elde edilmesi amaçlanmaktadır. Şekil 29'da çapraz doğrulama işleminin yapısı verilmiştir.



Şekil 29. Çapraz doğrulama işleminin yapısı

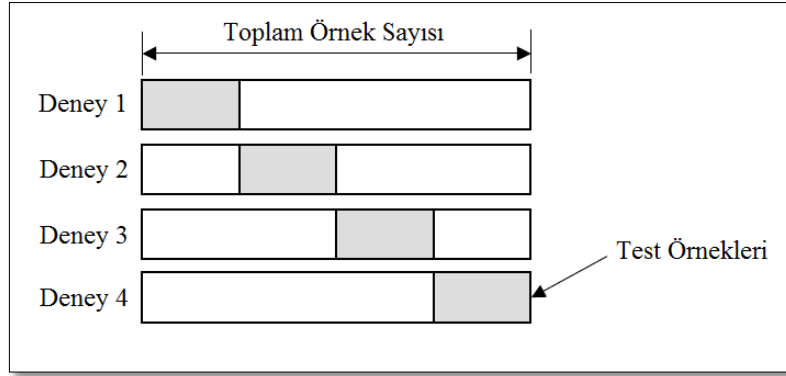
Literatürde mevcut olan üç temel doğrulama yöntemi vardır. Bunlar: Rastgele Alt Örnekleme, K-Kat Doğrulama, 'Birini Dışarıda Bırak' Doğrulama.

Rastgele alt örneklemede, eğitim verileri kendi içinde belli sayıda rastgele yerlerden seçilmiş örneklerle doğrulama ve ön eğitim olarak ayrılır.



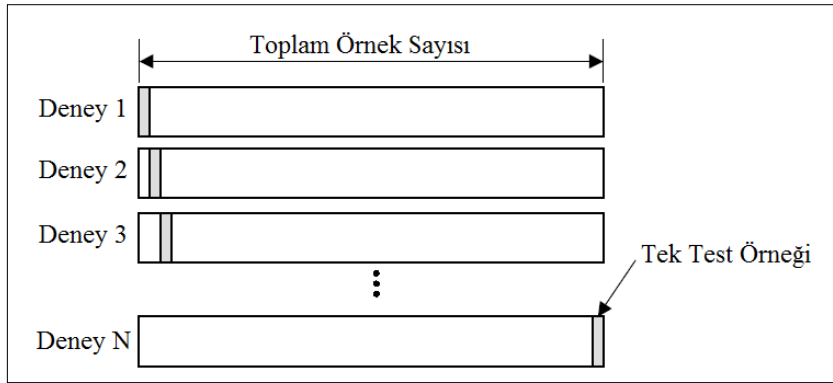
Şekil 30. Rastgele alt örnekleme (Random subsampling)

K-kat doğrulamada, veri kümesi eşit sayıda örneğe sahip, K tane alt örnek gruplarına ayrılır. Bu K tane alt örnek gruplarından bir tanesi doğrulama, K-1 tanesi ise ön eğitim verisi olarak belirlenir. Doğrulama süreci K kere tekrarlanır. K deney sonucunda elde edilen en yüksek doğrulukların bulunduğu durumlar üzerinden test işlemi yapılır.



Şekil 31. K-kat doğrulama

Birini dışarıda bırak (Leave-one-out) doğrulama, K-kat doğrulama algoritmasının özel bir halidir. K değeri burada örnek sayısına eşittir. Her bir örnek birer birer validasyona katılır. K-kat doğrulama algoritmasında iterasyon sayısını belirleyen K değeri olduğu için, burada da ne kadar örnek varsa, o kadar iterasyon gerçekleşecek ve o kadar sonuç bulunacaktır.



Şekil 32. 'Birini Dışarıda Bırak' doğrulama

2.5. Sınıflandırma Sonuçlarının Elde Edilmesi ve Karşılaştırılması

Bu bölümde, Bölüm 2.2.'de bahsedilen veri setinde mevcut olan üç farklı dönemin algılanması ve ayırt edilmesi işlemlerinden bahsedilecektir. Bu sebeple ilk olarak veri setinde mevcut olan kayıtlar, açıklamaları ve deneme sayıları Tablo 5'te verilmiştir.

Tablo 5. Kullanılan veri setleri ve açıklamaları

Sınıflar	Açıklama	Deneme Sayısı
A	Sağlıklı bireylerden göz açık durumda kaydedilen	100
B	Sağlıklı bireylerden göz kapalı durumda kaydedilen	100
C	Hastalıklı bireylerden nöbet olmayan durumda hippokampal oluşumdan kaydedilen	100
D	Hastalıklı bireylerden nöbet olmayan durumda kaydedilen	100
E	Hastalıklı bireylerden epileptik nöbet durumunda kaydedilen	100

Tablo 5'te belirtilen bu beş farklı sınıfın belli ve kurallı kombinasyonları elde edilerek alt veri kümeleri oluşturulmuş ve sınıflandırma işlemi bu alt veri kümeleri için ayrı ayrı gerçekleştirilerek performans hesaplanmıştır. Oluşturulan bu alt kümeler Tablo 4'te verilmiştir.

Alt kümelere ayırma işlemi, [17] çalışmasından esinlenilerek gerçekleştirilmiş; ancak bu çalışmadan farklı olarak, yalnızca epileptik nöbetin tespitine odaklanmak yerine, hastalıklı insanın nöbet geçirmediği evresinde kaydedilen ile sağlıklı insandan alınan işaretin ayrıştırılabilmesi için 7 numaralı küme oluşturulmuştur. Böylece epilepsi hastası bir insanın, nöbet geçirmediği durumda da sağlıklı insana göre farklılıklar taşıdığı fikri ortaya konulmaktadır.

Her bir alt küme için veriler %60 eğitim ve %40 test verisi olacak şekilde 10-kat çapraz doğrulama ile ayrılarak öznitelikler çıkarılmıştır. Çapraz doğrulamadan 2.4'te bahsedilmiştir. 7 tür öznitelik vektörü oluşturulmuş ve bunların her birisi için ayrı ayrı sınıflandırma performansı elde edilmiştir. Öznitelik vektörleri ve boyutları Tablo 6'da verilmiştir.

Tablo 6. Sınıflandırıcıya sunulacak olan öznitelik vektörleri ve boyutları

Vektör	Öznitelik Vektörü	Boyut
1	Hjorth Betimleyicileri	3
2	Entropi + Hjorth Betimleyicileri	4
3	TÖA Parametreleri	10
4	Hjorth Betimleyicileri + TÖA Parametreleri	13
5	Entropi + Hjorth Betimleyicileri + SGO + Hurst Üstel Katsayısı	6
6	Entropi + Hjorth Betimleyicileri + TÖA Parametreleri + SGO + Hurst Üstel Katsayısı	16
7	Entropi + SGO + Hurst Üstel Katsayısı	3

Öznitelik vektörleri çok katmanlı yapay sinir ağı, k-en yakın komşuluk, destek vektör makineleri ve Bayes sınıflandırıcılar kullanılarak sınıflandırılmıştır.

Çok katmanlı yapay sinir ağında, her bir vektör için ortak bir sistem oluşturacak şekilde parametreler ayarlanmıştır. Girişten verilen veriler minimum 3 ve maksimum 16 boyutlu olduğundan, ara katmanda kullanılacak olan nöron yapısı her bir öznitelik için çapraz doğrulama ile tespit edilmiş ve 40 olarak uygun görülmüştür. Öğrenme ve momentum katsayıları sırasıyla 0.1 ve 0.3 alınmıştır. Ağ eşiği olarak 0.001 alınmış ve aktivasyon fonksiyonu olarak sigmoid kullanılmıştır.

k-en yakın komşuluk sınıflandırıcısında, sınıflandırmayı etkileyen parametreler k (komşu sayısı) ve uzaklık ölçüm parametreleridir. k parametresinin belirlenmesinde çapraz doğrulama işlemi uygulanmıştır. Bu işlem, her bir öznitelik vektörü kümesine göre yeniden ayarlanarak sınıflandırma işlemi yapılmıştır. Bunun dışında, öznitelikler arasındaki uzaklığın ölçülmesinde ise Öklid uzaklığı kullanılmıştır.

Destek vektör makinesi sınıflandırıcısında, doğrusal olmayan ayrıştırma kullanılan çekirdek tabanlı destek vektör makinesi kullanılmıştır. Çekirdek fonksiyonu olarak radyal tabanlı çekirdek fonksiyonundan faydalanılmıştır. Bu fonksiyonda, radyal çapını belirleyen parametre γ , onu etkileyen parametre ise σ parametresidir. σ değerinin belirlenmesinde de

yine çapraz doğrulamaya başvurularak değer seçimi yapılmış ve 0.5 seçilerek sınıflandırma işlemi yapılmıştır.

Son olarak, Bayes sınıflandırıcıda her bir özneliği birbirinden bağımsız ele alan Naive Bayes sınıflandırıcı yöntemi kullanılmıştır. Bu yöntem olasılıksal dağılıma göre sınıflandırma yaptığından, özel bir parametre analizine gerek yoktur.

2.5.1. Sınıflandırıcı Performanslarının Elde Edilmesi

Bu bölümde, oluşturulan tüm öznelik vektörleri, her bir alt küme için mevcut sınıflandırıcıların her birine ayrı ayrı uygulanıp, her bir küme için en iyi doğruluğu veren öznelik vektörünün ve sınıflandırıcının tespit edilmesi amaçlanmıştır. Bunun için, ilk olarak öznelikler sabit tutulup, grupların performansı ölçülmüştür. Tablo 7'de Hjorth betimleyicilerinden oluşan 1 numaralı öznelik vektörüne ilişkin sonuçlar gösterilmiştir. Parantez içinde verilen değerler k-EYK için sırasıyla standart sapma ve en iyi k değeri olurken, diğer sınıflandırıcılar için yalnızca standart sapma değerleridir.

Tablo 7. 1. öznelik vektörüne ilişkin sınıflandırma performansları

Alt Veri Kümeleri	k-EYK	Bayes	DVM	YSA
Küme 1	61 (1.91, 19)	91 (1.54)	97 (0.99)	96 (1.22)
Küme 2	81,30 (2.00, 9)	95 (1.92)	91,60 (2.01)	94,16 (0.65)
Küme 3	73,75 (2.11, 11)	95 (1.83)	96,60 (1.72)	95,83 (1.34)
Küme 4	95,83 (0.68, 3)	98,30 (0.74)	96,60 (1.84)	96,60 (1.19)
Küme 5	97 (0.54, 5)	95,50 (0.62)	97,50 (1.32)	97,50 (1.49)
Küme 6	64,20 (2.23, 21)	93,50 (1.37)	93,50 (1.89)	97 (0.62)
Küme 7	60,68 (2.64, 25)	99,375 (0.98)	98,75 (1.11)	98,75 (0.66)

Burada, yalnızca Hjorth betimleyicileri kullanılmasına rağmen bazı kümelerin sınıflandırma performansları, başarılı sayılabilecek ölçüde yüksektir. Sınıflandırma başarısı en yüksek Bayes sınıflandırıcısı kullanılarak Küme 7 için ve %99.75, en düşükse k-EYK sınıflandırıcısı ile yine Küme 7 için %60,68 olarak bulunmuştur. Elde edilen bu değerler sınıflandırma doğruluğudur ve (65) bağıntısı ile hesaplanır.

$$\text{Doğruluk} = \frac{\text{Doğru_Pozitif} + \text{Doğru_Negatif}}{\text{Doğru_Pozitif} + \text{Yanlış_Pozitif} + \text{Doğru_Negatif} + \text{Yanlış_Negatif}} \quad (65)$$

Doğruluk performans ölçütü, pozitif değerlendirmelerin, tüm değerlendirmelere oranıdır. Buna göre;

Doğru_Pozitif : Pozitif kararın doğru olması olasılığını;

Doğru_Negatif : Negatif kararın doğru olması olasılığını;

Yanlış_Pozitif : Pozitif kararın yanlış olması olasılığını;

Yanlış_Negatif : Negatif kararın yanlış olması olasılığını göstermektedir.

Tablo 7'deki genel dağılıma bakıldığında; 2, 4 ve 7 nolu kümelerinin Bayes ile 1, 3 ve 5 kümelerinin DVM ile ve 5 ve 6 kümelerinin ise ÇKYSA ile en yüksek sınıflandırma doğruluklarını verdiklerini; k-en yakın komşuluk sınıflandırıcısının ise Küme 4 ve Küme 5 dışında başarı gösteremediği görülmektedir.

Hjorth betimleyicilerinden oluşan bu öznitelik vektörü için elde edilen sınıflandırma doğrulukları, geriye kalan 6 öznitelik vektörü için de aynı şekilde hesaplanmış ve her bir öznitelik vektörü için Tablo 7'deki gibi tablolar oluşturulmuştur. Bundaki amaç, her bir alt küme için en iyi sınıflandırmayı gerçekleştirecek olan öznitelik vektörünü ve sınıflandırıcıyı bulabilmektir.

Bir kümenin en iyi şekilde sınıflandırılması için bazen tek bir sınıflandırıcı kararı yeterli olmayabilir. Bu durumda sınıflandırıcıların melez kullanımı ile performans artırımı sağlanabilir. Literatürde bunu gerçekleyen birçok uygulama vardır. Örneğin [77] çalışmasında, melez aşırı öğrenme makinesi yöntemi kullanılmıştır. Bir diğer çalışmada, uyku durumu analizi amacıyla karar ağacı tabanlı YSA sınıflandırıcısı kullanılmıştır [78]. [79] çalışmasında negatif seçim ve kısmi yığın optimizasyonu tabanlı melez metot önerilmiştir. Son olarak [80] çalışmasında yine kısmi yığın optimizasyonu ve ANN yöntemlerinden elde edilen melez sınıflandırıcı yapısı ile EEG sinyalleri sınıflandırılmıştır.

Bu çalışmada ise diğer çalışmalardan farklı olarak, Borda sayısı metodu ile sınıflandırıcıların oylanması ve bu şekilde melez sınıflandırıcı sistemlerinin oluşturulması gerçekleştirilmiştir. Bu yöntemin probleme uyarlanışından bir sonraki bölümde bahsedilecektir.

2.5.2. Melez Yaklaşımlar ve Borda Sayısı ile Sınıflandırıcıların Ağırlıklandırılması

Borda sayısı, adayların önceliklerine göre seçim yapan bir oylama algoritmasıdır. Adaylar (çalışmada sınıflandırıcılara karşılık gelir), önceliklerine göre ödüllendirilir; bu toplam ödül puanına göre kazanan belirlenir. Genel olarak, N adaydan oluşan bir problem olduğunu varsayalım. İlk sırada oylanan aday N puan, ikinci sırada oylanan aday N-1 puan ve son olarak N. sırada oylanan aday 1 puan ile ödüllendirilir. Buradan en yüksek puanı toplayan aday kazanır [81].

P, Q, R ve S adayları oylanacak olsun. Oy verenler 5 alt gruba ayrılmış ve her bir alt grubun seçim öncelikleri tabloya yerleştirilmiştir. Örneğin 14 kişilik alt grubun oylama öncelikleri sırasıyla P, Q, R ve S şeklindedir. Bu durumda P'nin ödül puanı 4, Q'nun 3, R'nin 2 ve S'nin 1 puandır.

Tablo 8. Borda sayısı için örnek öncelik tablosu

Oylayanların Sayısı	14	10	8	4	1
Seçim 1	P	R	S	Q	R
Seçim 2	Q	Q	R	S	S
Seçim 3	R	S	Q	R	Q
Seçim 4	S	P	P	P	P

Tablo 8'deki tüm durumlar için ödül puanlar belirlendikten sonra, her bir adayın toplam puanın hesaplanması işlemi gerçekleştirilir. P adayının toplam puanı şu adımlarla hesaplanır: Birinci tercih için, oyunu P'den yana kullanan kişi sayısı 14 ve P'nin ödül puanı 4 olduğundan $14 \times 4 = 56$ bulunur. İkinci ve üçüncü tercihi P olan kişi sayısı 0'dır. Dördüncü yani son tercihi P olan kişi sayısı $10 + 8 + 4 + 1 = 23$ kişi ve bu seçimin ödül puanı 1 olduğundan, $23 \times 1 = 23$ bulunur. Toplam duruma bakıldığında $P = 56 + 23 = 79$ olarak sonuçlanır. Tüm adaylar için bu işlem yapıldığında; Q için 106, R için 104 ve S için 81 puan elde edilir. Sonuç olarak kazanan Q adaydır.

Sınıflandırma problemi açısından bu durumdan şu şekilde faydalanılmıştır. Tüm alt kümeler için her bir öznitelik vektörüne dair doğruluk değerleri çıkarılmış ve mevcut 4 adet sınıflandırıcının verdiği ortalama sınıflandırma doğruluğuna göre, öncelik tablosu oluşturulmuştur. Örneğin Tablo 7 için öncelik tablosu, Tablo 9'daki gibi oluşturulmuştur.

Burada P, k-EYK'yi; Q, Bayes sınıflandırıcıyı; R, destek vektör makinesini ve S, yapay sinir ağlarını ifade etmektedir.

Tablo 9. Hjorth betimleyicilerine göre oluşturulmuş öncelik tablosu

Seçimler	Küme 1	Küme 2	Küme 3	Küme 4	Küme 5	Küme 6	Küme 7
Seçim 1	R	Q	R	S	R	S	Q
Seçim 2	S	S	S	R	S	Q	R
Seçim 3	Q	R	Q	Q	P	R	S
Seçim 4	P	P	P	P	Q	P	P

Ancak burada Tablo 8'de olduğu gibi oylayan sayısı gibi bir kavram yoktur. Buradaki seçim, ağırlıklandırma amaçlı yapılacağı için başlangıçta değer her bir küme için sabit alınabilir. Bu çalışma için de bu değer her birisi için 1 alınmıştır. P, Q, R ve S değerleri yukarıdaki gibi hesaplanacak olursa;

$$P = 2 \times 1 + 1 \times 6 = 8$$

$$Q = 4 \times 2 + 3 \times 1 + 2 \times 3 + 1 \times 1 = 18$$

$$R = 4 \times 3 + 3 \times 2 + 2 \times 2 = 22$$

$$S = 4 \times 2 + 3 \times 4 + 2 \times 1 = 22$$

Tablodaki toplam puan 70'tir. P, Q, R ve S değerleri bu değerle oranlandığında her bir sınıflandırıcıya ait bir katsayı değeri elde edilecektir. Bu değer K ile gösterilecek olursa, aşağıdaki katsayılar elde edilmiş olur.

$$K_P = 8/70 = 0.115$$

$$K_Q = 18/70 = 0.257$$

$$K_R = 22/70 = 0.314$$

$$K_S = 22/70 = 0.314$$

Bu katsayılar sayesinde melez sınıflandırıcı sistemleri oluşturulurken, sınıflandırma performansları eşit tutulmamaktadır. Bundaki amaç, tek başına kötü performansa sahip olan bir sınıflandırıcı ile birleşen ve tek başına iyi performans gösteren sınıflandırıcının

başarısını kaybetmemesi ya da bir iki başarılı sınıflandırıcının birbirini olumlu yönde etkilemesi içindir.

2.5.3. Tekil ve Melez Sınıflandırma Sonuçları

Bu bölümde öncelikle her bir kümeye ait öznelik vektörlerinin sınıflandırma başarı performansından bahsedilip, daha sonra bu performansları iyileştirmeye yönelik olarak gerçekleştirilen melez yaklaşımdan bahsedilecektir.

Tekil sınıflandırıcı performansları hem öznelik vektörlerine hem de kümelere göre ayrı ayrı incelenmiştir. Her bir öznelik vektörü ve küme için oluşturulan bu durumların önce maksimum, daha sonra ortalama sonuçları verilmiştir. Buradan yola çıkılarak melez yaklaşımlara gidilmiş ve ne kadar artış olduğu ya da durumun korunup korunmadığı bilgisine ulaşılmıştır.

İlk olarak Tablo 10'da tüm alt kümeler için maksimum sınıflandırma doğrulukları ve maksimum sonucu veren öznelik vektörleri (parantez içinde) verilmiştir. Öznelik vektörleri ve açıklamaları daha önce Tablo 6'da verilmişti.

Tablo 10. Tüm alt kümeler için maksimum sınıflandırma doğrulukları

Veri Alt Kümeleri	Sınıflandırıcılar			
	k-EYK(%)	Bayes(%)	DVM(%)	YSA(%)
Küme 1	64,95 (7)	97 (5)	97 (1)	96.5 (4)
Küme 2	81,33 (2)	92.5 (4,6)	91.67 (1)	95.83 (4)
Küme 3	75.25 (6)	96.67 (2)	96.67 (1)	96.67 (6)
Küme 4	98.5 (5)	99.16 (5)	97.5 (3)	100 (5,6)
Küme 5	97.35 (7)	98 (2)	98 (2)	99,5 (7)
Küme 6	70.2 (4)	96.5 (4,6)	97 (2)	99.5 (6)
Küme 7	60.68 (1)	99.375 (1)	98.75 (1)	99.375 (6)

Bu tabloya göre, her bir kümenin başarılı olduğu farklı bir öznelik vektörü ve buna bağlı olarak farklı sınıflandırıcı performansları vardır. Buradan her bir küme için

sınıflandırma performanslarının k-EYK kullanıldığında en düşük, YSA kullanıldığında en yüksek performansı sağladığı görülmektedir.

Benzer şekilde Tablo 11'de, her bir öznitelik vektörü için maksimum performans değeri ve bunun gerçekleştiği veri kümesi (parantez içinde) verilmiştir.

Tablo 11. Her bir öznitelik vektörü için maksimum sınıflandırma doğrulukları

Öznitelik Vektörleri	Sınıflandırıcılar			
	k-EYK(%)	Bayes(%)	DVM(%)	YSA(%)
Öznitelik 1	97 (5)	99.375 (7)	98.75 (7)	98.75 (7)
Öznitelik 2	96.9 (5)	98 (5)	98 (5)	99 (5)
Öznitelik 3	95 (5)	98.33 (4)	97.5 (4)	99.16 (4)
Öznitelik 4	96.58 (4)	100 (4)	96.5 (5)	99.16 (4)
Öznitelik 5	98.5 (4)	99.16 (4)	96.88 (7)	100 (4)
Öznitelik 6	98.33 (4)	98.33 (4)	86.25 (7)	100 (4)
Öznitelik 7	97.35 (5)	97.5 (4,5)	96.5 (5)	99.5 (5)

Tablo 12'de her bir alt küme için öznitelik vektörleri ayrı ayrı değerlendirilmeksizin ortalama sınıflandırıcı başarısı incelenmiştir. Buradan 4 ve 5 numaralı kümelerin her sınıflandırıcı için yüksek sayılabilecek sınıflandırma başarılarının olduğu görülmektedir. Gruptan bağımsız olarak, özniteliklerin sınıflandırıcı performansları ise Tablo 13'te verilmiştir.

Tablo 12. Her bir alt küme için ortalama sınıflandırma doğrulukları

Veri Alt Kümeleri	Sınıflandırıcılar			
	k-EYK(%)	Bayes(%)	DVM(%)	YSA(%)
Küme 1	63,56	93	86,14	92,92
Küme 2	76,27	91,3	79,40	92,26
Küme 3	69,11	92,02	79,40	92,73
Küme 4	96,63	98,3	90,83	98,92
Küme 5	96	96,64	94,35	98,85
Küme 6	66,2529	90,92	89,286	96
Küme 7	56,55	94,01	91,60	95,98

Tablo 13. Her bir öznitelik vektörünün ortalama sınıflandırıcı performansları

Öznitelik Vektörleri	Sınıflandırıcılar			
	k-EYK(%)	Bayes(%)	DVM(%)	YSA(%)
Öznitelik 1	76,25	95,38	95,96	96,55
Öznitelik 2	75,29	95,17	95,53	96,95
Öznitelik 3	73,84	93,37	86,82	93,71
Öznitelik 4	73,73	95,52	84,69	96,92
Öznitelik 5	75,41	96,04	85,95	96,68
Öznitelik 6	75,32	95,65	73,53	97,17
Öznitelik 7	74,52	85,09	88,80	89,69

Melez yaklaşım sayesinde iki veya daha fazla sınıflandırıcı beraber kullanılarak, ortalama sınıflandırıcı performansının artırılması amaçlanmış ve başarı sağlanmıştır. Bölüm 2.5.2’de bahsedildiği üzere, her bir öznitelik vektörüne göre elde edilen ortalama dağılımdan Borda öncelik tablosu oluşturulmuş ve bu sayede sınıflandırıcıların etkisi özniteliklere göre dinamik bir şekilde ayarlanmıştır.

10 adet melez sınıflandırıcı sistemi oluşturulmuştur. Bunlar Tablo 14’te ifade edilmiştir. Her bir öznitelik için ağırlıklandırmalar yapılmış ve buna göre her bir küme Tablo 14’teki her bir sınıflandırıcı sistemine uygulanarak, sınıflandırma doğruluğu ölçülmüştür. Buna ilişkin bilgiler ve başarıyı sağlayan öznitelik vektörü (parantez içinde) ise Tablo 15’te gösterilmiştir. Buna göre sırasıyla Küme 1 için DVM ve YSA; Küme 2 için k-EYK, Bayes ve YSA; Küme 3 için Bayes, DVM ve YSA; Küme 4 için 2 ve 3 numaralı sistemler hariç tümü; Küme 5 için k-EYK ve YSA; Küme 6 için 4, 7 ve 10 numaralı sistemler ve son olarak Küme 7 için 5 ve 6 numaralı sistemler dışındaki tüm sistemler yüksek sınıflandırma performansı sağlamışlardır. Sınıflandırma sistemlerinin kendi içerisinde ortalama sınıflandırma performanslarına bakılacak olursa, DVM ve YSA’dan oluşan sınıflandırma sisteminin başarısının diğerlerine oranla daha yüksek olduğu sonucuna varılmıştır.

Tablo 14. Kullanılan melez sınıflandırıcı sistemleri

Sınıflandırıcı Sistemi	Sınıflandırıcılar
S1	k-EYK + Bayes + DVM
S2	k-EYK + Bayes + YSA
S3	Bayes + DVM + YSA
S4	k-EYK + DVM + YSA
S5	k-EYK + Bayes
S6	k-EYK + DVM
S7	k-EYK + YSA
S8	Bayes + DVM
S9	Bayes + YSA
S10	DVM + YSA

Tablo 15. Melez sınıflandırıcı sistemlerine ilişkin sınıflandırma performansları

Alt Veri Kümeleri	Melez Sınıflandırıcı Sistemleri									
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Küme 1	97	97	95	96	97	96	75	94,5	93,5	97,5 (1)
Küme 2	94.16	97,5 (5)	94,16	95.83	94,16	95.83	93.33	92.5	95.83	93.33
Küme 3	93.33	96.66	98.33 (1)	96.66	94.16	80	96.66	96.66	97,5	96.66
Küme 4	100 (4-5)	99.16	99.16	100 (5-6)	100 (4)	100 (5)	100 (5-6)	100 (4)	100 (4-5)	100 (5-6)
Küme 5	98	99	99	99	98	98	99.5 (7)	98	99	99
Küme 6	96.5	99	99	99.5 (6)	86.5	86.5	99.5 (6)	96.5	99	99.5 (6)
Küme 7	99.375 (1-3)	99.375 (1-3)	99.375 (1)	99.375 (6)	80.625	81.25	99.38 (6)	99.375 (1)	99.375 (1)	99.375 (6)
Ortalama	96,90	98,24	97,47	98,28	92,92	90,29	94,45	96,79	98,06	99,07

Tablo 15'te görülen maksimum sınıflandırma doğruluğu hesabının yanı sıra, diğer iki performans kriteri de duyarlılık ve özgünlüktür. Bunların hesaplanması (66) ve (67) eşitliklerinde ifade edilmiştir.

$$\text{Duyarlılık} = \frac{\text{Doğru_Pozitif}}{\text{Doğru_Pozitif} + \text{Yanlış_Negatif}} \quad (66)$$

$$\text{Özgünlük} = \frac{\text{Doğru_Negatif}}{\text{Doğru_Negatif} + \text{Yanlış_Pozitif}} \quad (67)$$

Buna göre Tablo 15'te verilen maksimum doğruluklara sahip sınıflandırıcı sistemlerine ilişkin duyarlılık ve özgünlük değerleri de Tablo 16'da verilmiştir.

Tablo 16. Sınıflandırma sonuçlarına ilişkin duyarlılık, özgünlük ve doğruluk oranları

Küme	Öznitelik	Sınıflandırıcı	Duyarlılık (%)	Özgünlük (%)	Doğruluk (%)
Küme 1	Hjorth betimleyicileri	DVM YSA	100	98,36	97,5
Küme 2	Hjorth betimleyicileri Shannon Entropi Hurst üsteli SGO	k-EYK Naive Bayes YSA	97,5	98,75	97,5
Küme 3	Hjorth betimleyicileri	DVM YSA Naive Bayes	100	97,56	98,33
Küme 4	Hjorth betimleyicileri TÖA	DVM YSA Naive Bayes	100	100	100
Küme 5	Shannon entropi Hurst üsteli SGO	k-EYK YSA	99,37	100	99,5
Küme 6	TÖA Hjorth betimleyicileri Hurst Üsteli Shannon entropi SGO	DVM YSA k-EYK	100	99,17	99,5
Küme 7	TÖA Hjorth betimleyicileri Hurst Üsteli Shannon entropi SGO	DVM YSA Naive Bayes k-EYK	98,76	100	99,38

Bir sonraki bölümde, elde edilen performanslar, literatürde bu çalışmada kullanılan veri setlerinden faydalanan diğer çalışmalarla karşılaştırılacaktır.

2.5.4. Sonuçların Karşılaştırılması

Literatürde bu çalışmada kullanılan veri setini kullanan ve oldukça yüksek sınıflandırma doğruluğu elde eden çalışmalar mevcuttur. Bu çalışmada olduğu gibi, bu verilerin kullanıldığı diğer çalışmalarda da veriler birtakım alt kümelere ayrılmış ve bu alt kümelerin ayrı ayrı sınıflandırılması işlemi gerçekleştirilmiştir. Bu tez çalışmasında oluşturulan veri gruplarının sınıflandırma performansları ve literatürdeki yerleri sırasıyla verilecektir.

İlk olarak normal, nöbet öncesi ve nöbet dönemlerini ayırt eden çalışmalardaki ilk 5 sınıflandırma doğruluğu verilmiştir. Buna göre tez çalışması 4. sırada yer almıştır.

Tablo 17. Sağlıklı, nöbet öncesi ve nöbet dönemlerine ilişkin performans tablosu

Sıralama	Yazarlar	Yılı	Öznitelik Çıkarma Tekniği	Sınıflandırma Metodu	Doğruluk (%)
1	Martinez-Vargas vd.	2011	Zaman frekans analizi	Eşikleme fonksiyonu	99
2	Guo vd.	2010b	Dalgacık dönüşümü Çizgi uzunluğu özelliği	Yapay sinir ağları	97.77
3	Tzallas vd.	2007a	Zaman-frekans analizi	Yapay sinir ağları	97.72
4	Tez çalışması		Hjorth betimleyicileri	DVM YSA	97.5
5	Orhan vd.	2011	Dalgacık Dönüşümü	k-EYK YSA	95.60

İkinci olarak ise göz açık durumda kaydedilen sağlıklı, nöbet öncesi ve nöbet dönemlerinin ayrılmasına yönelik ilk çalışma verilmiştir.

Tablo 18. Sağlıklı insanda göz açık durumda kaydedilen veriler ile nöbet öncesi ve nöbet durumlarının sınıflandırma performansları

Sıralama	Yazarlar	Yılı	Öznitelik Çıkarma Tekniği	Sınıflandırma Metodu	Doğruluk (%)
1	Tzallas	2009	Zaman-frekans analizi	Naive Bayes k-en yakın komşuluk Karar ağacı Yapay sinir ağları	100
	Tzallas vd.	2007b	Zaman-frekans analizi	Yapay sinir ağları	100
2	Dastidar vd.	2008	Temel bileşenler analizi	Yapay sinir ağları	99.3
3	Tzallas vd.	2007a	Zaman-frekans analizi	Yapay sinir ağları	99.28
4	Liang vd.	2010	Zaman-frekans analizi Yaklaşık entropi	Doğrusal en küçük kareler Doğrusal ayırtaç Destek vektör makineleri	98.67
5	Tez çalışması		Hjorth betimleyicileri Shannon Entropi Hurst üsteli, SGO	k-EYK Naive Bayes YSA	97.5

Üçüncü durum olan, göz kapalı halde alınan sağlıklı kaydın ve nöbet öncesi ve nöbet dönemlerinin sınıflandırılmasına ilişkin literatür çalışması yoktur. Dolayısıyla şu an için en yüksek doğruluk bu tez çalışmasında elde edilmiş ve %98.33'tür. Bu doğruluğun elde edilmesinde öznitelik olarak Hjorth betimleyicilerinden ve sınıflandırıcı sistemi olarak Naive Bayes, SVM ve YSA sınıflandırıcılarından faydalanılmıştır.

Dördüncü sınıflandırma işlemi, sağlıklı dönem (göz açık ve kapalı) ile yalnızca nöbet döneminin ayırt edilmesine yöneliktir ve bununla ilgili de literatür çalışmasına rastlanamamıştır. Melez sistemin en istikrarlı çalıştığı veri grubudur çünkü, 2 ve 3 numaralı sınıflandırma sistemleri hariç tüm durumlarda %100 başarı sağlamıştır.

Beşinci veri grubu, nöbet döneminin diğer tüm dönemlerden ayrıldığı durumdur. Bu küme, sadece nöbet verisinin ayırt ediciliğini ölçmek amacıyla oluşturulmuştur. Buna yönelik çalışma performansları da tabloda verilmiştir.

Tablo 19. Nöbet verisinin diğer durumlardan ayırt edilme performansı

Sıralama	Yazarlar	Yılı	Öznitelik Çıkarma Tekniği	Sınıflandırma Metodu	Doğruluk (%)
1	Orhan vd.	2011	Dalgacık Dönüşümü	k-EYK YSA	100
2	Tez çalışması		Shannon entropi Hurst üsteli SGO	k-EYK YSA	97.5
3	Liang vd.	2010	Zaman-frekans analizi Yaklaşık entropi	Doğrusal en küçük kareler Doğrusal ayırtaç Destek vektör makineleri	98.51
4	Guo vd.	2010a	Dalgacık dönüşümü Yaklaşık Entropi	Yapay sinir ağları	98.27
5	Guo vd.	2010b	Dalgacık dönüşümü Çizgi uzunluğu özelliği	Yapay sinir ağları	97.75

Altıncı veri grubu ise sağlıklı ve epilepsi hastası bireyleri genel olarak ayıran veri kümesidir ve başarı sıralaması tabloda verilmiştir. Bu küme için literatür çalışmaları içindeki en yüksek sınıflandırma performansı elde edilmiştir.

Tablo 20. Sağlıklı ve epilepsi hastası bireylerin genel olarak ayırt edilme performansları

Sıralama	Yazarlar	Yılı	Öznitelik Çıkarma Tekniği	Sınıflandırma Metodu	Doğruluk (%)
1	Tez çalışması		TÖA Hjorth betimleyicileri Hurst Üsteli Shannon entropi SGO	DVM YSA k-EYK	99.5
2	Orhan vd.	2011	Dalgacık Dönüşümü	k-EYK YSA	98.80
3	Hekim	2013	Entropi	YSA	96.80

Son olarak nöbet anında olmayan epilepsi hastası bir insan ile sağlıklı bir insanı sınıflandırmak amacıyla oluşturulan yedinci veri kümesinin sınıflandırma performansı verilmiştir. Bu durumda da şu an için çalışma en yüksek sınıflandırma performansını vermektedir.

Tablo 21. Sağlıklı ve nöbet olmayan durumdaki verilerin ayırt edilme performansları

Sıralama	Yazarlar	Yılı	Öznitelik Çıkarma Tekniği	Sınıflandırma Metodu	Doğruluk (%)
1	Tez çalışması		TÖA Hjorth betimleyicileri Hurst Üsteli Shannon entropi SGO	DVM YSA Naive Bayes k-EYK	99.375
2	Hekim	2013	Entropi	YSA	96.20

3. SONUÇLAR VE ÖNERİLER

Bu tez kapsamında, epilepsi hastası bir bireyin sağlıklı bireylerden ya da epilepsi nöbeti geçirmediği durumdan ayrılmasına yönelik otomatik teşhis yapan bir sistem gerçekleştirilmiştir. Normalde gözle görülemeyecek ayrıntıların, matematiksel modellemelerle bilgisayarlı sistemlere öğretilip, genel ve objektif tanıma yapılması amaçlanmıştır.

Bunun için çalışma çapraz doğrulama ile eğitim ve test veri setlerinin oluşturulması, öznitelik çıkarımı ve sınıflandırma olmak üzere üç temel kısım üzerine kurulmuştur. Toplamda 5 farklı gruptan alınan ve her birinde 100 deneme bulunan veri seti mevcuttur ve bu veri setinde sağlıklı, nöbet öncesi ve nöbet olmak üzere 3 farklı sınıf bilgisi vardır. Amacına göre veri kümesinin belirli seçimleriyle 7 alt küme oluşturulmuş ve bu şekilde 7 adet durum değerlendirilmesi yapılmıştır.

Eğitim ve test verileri üzerinden zaman uzayında öznitelik çıkarımı gerçekleştirilmiştir. Özniteliklerde belli seçimlerle 7 sınıfa ayrılmış ve sınıflandırma değerlendirmesi her bir durum için ayrı yapılmıştır. Sınıflandırıcı olarak 4 adet sınıflandırıcı kullanılmış, bu sınıflandırıcılar önce tekil daha sonra melez yakшалımlarla değerlendirilmiştir.

Sınıflandırıcı performanslarının artırılması için, eğitim verileri ön eğitim ve doğrulama verileri olarak ayrılmış ve çapraz doğrulama yöntemleriyle sınıflandırıcı parametreleri seçilmiştir. k -EYK için k değeri, YSA için ara katman ve ara katmandaki nöron sayıları ve DVM için radyal tabanlı çekirdek fonksiyonun genişliği parametrelerinin optimum değerleri 10-kat çapraz doğrulama yöntemi sayesinde her bir sınıflandırma işlemi için elde edilmiştir.

Tekil olarak gerçekleştirilen sınıflandırma işlemi hem oluşturulan alt kümelere, hem de öznitelik vektör kümelerine göre ayrı ayrı gerçekleştirilmiştir. Elde edilen ortalama sınıflandırma doğrulukları bulunmuştur. Melez yöntemlere geçişte de bu dağılıma bakılarak sınıflandırma ağırlıklarını hesaplamak ve bu sayede yüksek doğruluk elde etmek amaçlanmıştır. Çünkü yapılan tekil sınıflandırmada her bir sınıflandırıcının aynı performansta olmadığı ve dolayısıyla melez bir sistemde aynı ağırlıkta olmamaları

gerektiđi düşünölmüştür. Ađırlıklandırmanın hesaplanmasına, oylama yöntemlerinden birisi olan Borda sayısı tekniđi probleme uyarlanarak çözümler getirilmiştir.

Melez sistemlerin oluşturulmasıyla elde edilen sınıflandırıcı sistemleri, beklenildiđi gibi tekil sınıflandırıcı performanslarına göre yüksek sınıflandırma doğruluđu vermişler ve sistem performansını artırmışlardır.

Literatür değerlendirilmesi yapıldığında aşğıdaki üç durum için en yüksek performansın bu tez çalışmasında elde edildiđi görölmüştür:

- Sağlıklı insanda göz kapalı durumda elde edilen verinin nöbet öncesi ve nöbet dönemlerinden ayırt edilmesi
- Sağlıklı ve epilepsi hastası (nöbet ve nöbet öncesi) bireylerin genel olarak sınıflandırılması
- Sağlıklı ve epilepsi hastası bireyin nöbet öncesi dönemlerinin ayırt edilmesi

Öneriler olarak; epilepsi hastalığı nöbetli bir hastalık olduđu için çalışmada bir deney ortamında verilerin oluşturulması mümkün olmamaktadır. Bu sebeple herkese açık bir veri seti kullanılmıştır. Yine bu şekilde herkese açık olan veri setlerine erişilerek ya da tıp doktorlarından veriler temin edilerek veri tabanı genişletilebilir.

Hedeflenen en yüksek başarıya ulaşmak için daha farklı öznelik çıkarma (örneğin, frekans uzayı yöntemleri) ve sınıflandırma yöntemleri denenebilir.

Veriler tıp doktorları ile birlikte yorumlanarak olası başka durumların da ortaya çıkarılması ve gerçek hayatta kullanıma uygun hale getirilmesi için adımlar atılabilir.

Çalışma sırasında uygulanan yaklaşımlara benzer yaklaşımlar sıçanlardan alınan ECoG verileri üzerinde uygulanarak bu üç dönemin ayırt edilmesi gerçekleşmiş; 2013 Sinyal İşleme Kurultayı ve Çankaya Üniversitesi 6. Mühendislik ve Teknoloji Sempozyumu'na bu çalışmaya ilişkin 2 adet bildiri sunulmuş ve bildiri kitapçıklarında basılmıştır [82] [83].

4. KAYNAKLAR

1. Aydemir, Ö., İmlecın İki Boyutlu Hareketinin Hayali Sırasında Kaydedilmiş EEG İşaretlerinin Karar Ağaç Yapısı Esaslı Sınıflandırılması, Doktora Tezi, K.T.Ü., Fen Bilimleri Enstitüsü, Trabzon, 2013.
2. Reynolds, P. E. H., Epilepsy: The Disorder, Epilepsy Atlas, (2005) 15-27.
3. Srinivasan, R., Methods to Improve the Spatial Resolution of EEG, International Journal of Bioelectromagnetism (IJBEM), 1,1 (1999) 102-111.
4. Subasi, A. ve Erçelebi, E., Classification of EEG Signals Using Neural Network and Logistic Regression, Computer Methods and Programs in Biomedicine, (2005) 87-99.
5. Ansari-Asl, K., Chanel, G. ve Pun, T., A Channel Selection Method for EEG Classification in Emotion Assessment Based on Synchronization Likelihood, 15th European Signal Processing Conference (EUSIPCO), 2007, Poland, Bildiri Kitabı, 1241-1245.
6. Cecchin, T., Ranta, R., Koessle, L., Caspary, O., Vespignani, H. ve Maillard, L., Seizure Lateralization in Scalp EEG Using Hjorth Parameters, Clinical Neurophysiology, 121 (2010) 290–300.
7. Orhan, U., Hekim, M. ve Ozer, M., EEG Signals Classification Using the K-means Clustering and A Multilayer Perceptron Neural Network Model, Expert Systems with Applications, 38 (2011) 13475–13481.
8. Ocak, H., Automatic Detection of Epileptic Seizures in EEG Using Discrete Wavelet Transform and Approximate Entropy, Expert Systems with Applications, 36 (2009) 2027-2036.
9. Subasi, A., Application of Adaptive Neuro-fuzzy Inference System for Epileptic Seizure Detection Using Wavelet Feature Extraction, Computers in Biology and Medicine, 37 (2007) 227-244.
10. Andrzejak R. G., Lehnertz, Mormann, F., Rieke, David, P. ve Elger, C.E., Indications of Nonlinear Deterministic and Finite Dimensional Structures in Time Series of Brain Electrical Activity: Dependence on Recording Region and Brain State, Phys. Rev. E, 64 (2001) 061907-1,8.
11. Gautama, T., Mandic, D. P. ve Hulle, M. M. V., Indications of Nonlinear Structures in Brain Electrical Activity, Phys. Rev. E, 67,4 (2003) 046204-1,5.
12. Kannathal, N., Acharya, U. R., Limb, C. M. ve Sadasivan, P. K., Characterization of EEG - A Comparative Study, Computer Methods and Programs in Biomedicine, 80, 1 (2005) 17-23.

13. Kannathal, N., Choo, M. L., Acharya, U. R. ve Sadasivan, P. K., Entropies for Detection of Epilepsy in EEG, Computer Methods and Programs in Biomedicine, 80, 3 (2005) 187-194.
14. Adeli, H., Ghost-Dastidar, S. ve Dadmehr, N., A Wavelet-Chaos Methodology for Analysis of EEGs and EEG Subbands to Detect Seizure and Epilepsy, IEEE Trans Biomed Eng., 54 (2007) 205-211.
15. Sadati, N., Mohseni, H. R. ve Maghsoudi, A., Epileptic Seizure Detection Using Neural Fuzzy Networks, 2006 IEEE International Conference on Fuzzy Systems, July 2006, Canada, Bildiri Kitabı: 596-600.
16. Srinivasan, V., Eswaran, C. ve Sriraam, N., Approximate Entropy-Based Epileptic EEG Detection Using Artificial Neural Networks, IEEE Transactions on Information Technology in Biomedicine, 11, 3 (2007) 288-295.
17. Tzallas, A. T., Tsipouras, M. G. ve Fotiadis, D. I., Automatic Seizure Detection Based on Time-Frequency Analysis and Artificial Neural Networks, Computational Intelligence and Neuroscience, 2007 (2007) 80510-1,13.
18. Polat, K. ve Güneş, S., Classification of Epileptiform EEG Using a Hybrid System Based on Decision Tree Classifier and Fast Fourier Transform, Applied Mathematics and Computation, 187, 2 (2007) 1017-1026.
19. Martinez-Vargas, J. D., Avendano-Valencia, L. D., Giraldo, E. ve Castellanos-Dominguez, G., Comparative Analysis of Time Frequency Representations for Discrimination of Epileptic Activity in EEG Signals, 5th International IEEE EMBS Conference on Neural Engineering, April 2011, Cancun, Bildiri Kitabı: 148-151.
20. Guo, L., Rivero, D., Dorado, J., Rabunal, J. R. ve Pazos, A., Automatic Epileptic Seizure Detection in EEGs Based on Line Length Feature and Artificial Neural Networks, J NeurosciMethods, 191,1 (2010) 101-109.
21. Ghosh-Dastidar, S., Adeli, H. ve Dadmehr, N., Mixed-Band Wavelet-Chaos-Neural Network Methodology for Epilepsy and Epileptic Seizure Detection, IEEE Trans Biomed Eng., 54, 9 (2007) 1545-1551.
22. Liang, S. F., Wang, H. C., Chang, W. L., Combination of EEG Complexity and Spectral Analysis for Epilepsy Diagnosis and Seizure Detection, EURASIP Journal on Advances in Signal Processing, 2010 (2010) 853434.
23. Cain, L., Brain and Spine: The Body's Control Center, <http://www.mtnsf.org> 15 Kasım 2013
24. Bilim, M., Beynin yapısı ve işleyişi, <http://aileportal.com/beynin-yapisi-ve-isleyisi.html> 15 Kasım 2013
25. Huguenard, J. ve McCormick, D. A., Electrophysiology of the Neuron, Oxford University Press, Newyork , 1994.

26. Giger, M. L. and Karssemeijer, N., Progress in Biomedical Optics and Imaging, Medical Imaging, 9, 32 (2008).
27. Aydemir, Ö. ve Kayıkçıoğlu, T., Dalgacık Dönüşümü Öznelikleri Kullanarak Motor İmgeleme ve ECoG İşaretlerinin Sınıflandırılması, IEEE 18. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, 2010, Diyarbakır.
28. <http://www.eeguzerine.com>, EEG Elektroensefalografi Nedir?. 15 Kasım 2013.
29. Rangayyan, R. M., Biomedical Signal Analysis, Jonh Wiley & Sons Inc., Canada, 2002.
30. Aydemir, Ö., Beyin Bilgisayar Arayüzü Uygulamalarına Yönelik EEG İşaretleri İçin Öznelik Çıkarma, Yüksek Lisans Tezi, K. T. Ü., Fen Bilimleri Enstitüsü, Trabzon, 2008.
31. Graimann, B., Allison, B. ve Pfurtscheller, G., Brain-Computer Interfaces, Springer, Newyork, 2010.
32. Yeni, S. N., Epilepsi ve Acil Sorunlar, İ.Ü. Cerrahpaşa Tıp Fakültesi Sürekli Tıp Eğitimi Etkinlikleri, 29 (2002) 219-236.
33. Yeni, S. N. ve Karaağaç, N., Epilepsiler, İ.Ü. Cerrahpaşa Tıp Fakültesi Sürekli Tıp Eğitimi Etkinlikleri - Nörolog Olmayanlar İçin Nöroloji Sempozyum Dizisi, 23 (2005) 159-175.
34. <http://www.turkepilepsi.org> Türk Epilepsi ile Savaş Derneği. 15 Kasım 2013
35. Hjorth, B., Time Domain Descriptors and Their Relation to a Particular Model for Generation of EEG Activity, CEAN – Computerized EEG Analysis, (1975) 3–8.
36. Gudmundsson, S., Runarsson, T. P., Sigurdsson, S., Eiriksdottir, G. ve Johnsen, K., Reliability of Quantitative EEG Features, Clinical Neurophysiology, 118 (2007) 2162–2171.
37. Eckmann, J. P., Kamphorst, S. O. ve Ruelle, D., Recurrence Plots of Dynamical Systems, Europhy. Lett., 4 (1987) 973-977.
38. Acharya, U. R., Sree, S. V., Swapna, G., Martis, R. J. ve Suri, J. S., Automated EEG Analysis of Epilepsy: A Review, Knowledge-Based Systems, 45 (2013) 147-165.
39. Acharya, U. R., Sree, S. V., Chattopadhyay, S., Wenwei, Y. U. ve Alvin, A. P. C., Application of Recurrence Quantification Analysis for the Automated Identification of Epileptic EEG Signals, Int. J. Neural Syst. 21,3 (2011) 199-211.
40. Chen, C. H., Signal Processing Handbook, Dekker, Newyork, 1988.

41. Staudinger, T. ve Polikar, R., Analysis of Complexity Based EEG Features for the Diagnosis of Alzheimer's Disease, 33rd Annual International Conference of the IEEE EMBS, August 2011, Boston, Bildiri Kitabı: 2033-2036
42. Osorio, I. ve Frei, M.G., Hurst Parameter Estimation for Epileptic Seizure Detection, Communications in Information and Systems, 7, 2 (2007) 167-176.
43. Najarian, K. ve Splinter, R., Biomedical Signal and Image Processing, CRC Press, United States, 2012.
44. Aydın, S., Saraoğlu, H. M. ve Kara, S., Epilepside EEG Tabanlı Entropi Değişimleri VI. Ulusal Tıp Bilişimi Kongresi Bildirileri, TURKMIA'09 Proceedings, 2009, İstanbul.
45. Rojas, R., Neural Networks, Springer-Verlag, Berlin, 1996.
46. Yegnanarayana, B., Artificial Neural Networks, Prentice Hall of India, New Delhi, 2006.
47. Öztemel, E., Yapay Sinir Ağları, Papatya Yayıncılık, İstanbul, 2006.
48. Bayır, F., Yapay Sinir Ağları ve Tahmin Modellemesi Üzerine Bir Uygulama, Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul, 2006.
49. Larose, D. T., Discovering Knowledge in Data, A John Wiley & Sons Inc. Publication, Newyork, 2005.
50. Olgun, M. O. ve Özdemir, G., İstatistiksel Özellik Temelli Bayes Sınıflandırıcı Kullanarak Kontrol Grafiklerinde Örüntü Tanıma, Journal of the Faculty of Engineering & Architecture of Gazi University, 27, 2 (2012) 303.
51. Jaynes, E. T., Probability Theory-The Logic of Science, Cambridge University Press, UK, 2003.
52. Squire, D. M., CSE5230 Tutorial: The Naive Bayes Classifier, Monash University, Faculty of Information Technology, Malaysia, 2004.
53. Liu, D. Y., Bayesian Classification, Wright State University, USA, 2009.
54. Chen, D., Sisterhood of Classifiers: A Comparative Study of Naive Bayes and Noisy-or Networks, University of California, Loas Angeles, CA.
55. Vapnik, V., The Nature of Stastical Learning Theory, Springer-Verlag, Newyork, 1995.
56. Vapnik, V., Stastical Learning Theory, John Wiley & Sons. Inc, Newyork, 1998.
57. Abe, S., Support Vector Machines for Pattern Classification, Springer, Newyork, 2010.

58. Aykut, M., Örüntü Tanıma Amaçlı Doğrusal Olmayan Yöntemlerin Geliştirilmesi ve Uygulanması, Yüksek Lisans Tezi, K. T. Ü., Fen Bilimleri Enstitüsü, Trabzon, 2007.
59. Cortes, C. ve Vapnik, V., Support Vector Networks, Machine Learning, 20 (1995) 273-297.
60. Cristianini, N. ve Shawe-Taylor, J., An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods, Cambridge University Press, UK, 2000.
61. Garg, G., Sing, V., Grover, M., Nidhi ve Gupta J. R. P., Optimal Kernel Learning for EEG Based Sleep Scoring System, International Journal of Biological & Medical Research, 2, 4 (2011) 1220-1225.
62. Lima C. A., Kernel Machines for Epilepsy Diagnosis via EEG Signal Classification: a Comparative Study, Artif Intell Med., 53, 2 (2011) 83-95.
63. Subasi, A. ve Gursoy, M. I., EEG Signal Classification Using PCA, ICA, LDA and Support Vector Machines, Expert Systems with Applications, 37,12 (2010) 8659-8666.
64. Pal, M., Multiclass Approaches for Support Vector Machine Based Land Cover Classification, National Institute of Technology, India, 2008.
65. Kugler, M., Divide-and-Conquer Large-Scale Support Vector Classification, Nagoya Institute of Technology, Japan, 2006.
66. Andrzejak, R. G., Klinik für Epileptologie, University of Bonn, <http://www.meb.uni-bonn.de/epileptologie/science/physik/eeodata.html> 15 Kasım 2013
67. Vidaurre, C., Kramer, N., Blankertz, B. and Schlögl, A., Time Domain Parameters as a Feature for EEG-Based Brain-Computer Interfaces, Neural Networks, 22 (2009) 1313-1319.
68. Zbilut, J. P. ve Webber, J. C. L., Recurrence Quantification Analysis: Introduction and Historical Context, International Journal of Bifurcation and Chaos, 17, 10 (2007).
69. Zhu, T., Huang, L., Zhang, S. ve Huang, Y., Predicting Epileptic Seizure by Recurrence Quantification Analysis of Single-Channel EEG, ICIC, 2008, Shanghai, China.
70. Ouyang, G., Xie, L., Chen, H., Li, X., Guan, X. ve Wu, H., Automated Prediction of Epileptic Seizures in Rats with Recurrence Quantification Analysis, Engineering in Medicine and Biology 27th Annual Conference, January 2005, Shanghai, China.
71. Li, X., Ouyang, G., Yao, X. ve Guan, X., Dynamical Characteristics of Pre-epileptic Seizures in Rats with Recurrence Quantification Analysis, Physics Letters A, 33 (2004) 164-171.

72. Goshvarpour, A. ve Goshvarpour, A., Recurrence Plots of Heart Rate Signals during Meditation, International Journal of Image, Graphics and Signal Processing, 2 (2012) 44-50.
73. Kurths, J., Marwan, N. ve Wessel, N., Recurrence Plot Based Measures of Complexity to Predict Life-Threatening Cardiac Arrhythmias, Proceedings ECCTD 03, 2003, Krakow.
74. Zandi, A. S., Tafreshi, R., Javidan, M. ve Dumont, G. A., Predicting Temporal Lobe Epileptic Seizures Based on Zero-Crossing Interval Analysis in Scalp EEG, 32nd Annual International Conference of the IEEE EMBS, August 2010, Buenos Aires, Argentina.
75. Conradsen, I., Beniczky, S., Hoppe, K., Wolf, P. ve Sorensen, H. B. D., Automated Algorithm for Generalized Tonic-Clonic Epileptic Seizure Onset Detection Based on sEMG Zero-Crossing Rate, IEEE Transactions on Biomedical Engineering, 59, 2 (2012) 579-585.
76. Shenga, H. ve Chen, Y., Multifractional Property Analysis of Human Sleep EEG Signals, Proceedings of the ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, 2011, Washington, DC, USA.
77. Geetha, S. G., Detecting Epileptic Seizures Using Electroencephalogram: A New and Optimized Method for Seizure Classification Using Hybrid Extreme Learning Machine, International Conference on Process Automation, Control and Computing (PACC), July 2011, Coimbatore.
78. Haoyu, M., Hu, B., Jackson, M., Jingzhi, Y. ve Wen, Z., A Hybrid Classification Method using Artificial Neural Network Based Decision Tree for Automatic Sleep Scoring, World Academy of Science, Engineering and Technology, 55 (2011) 279.
79. Ba-Karait, N. O., Shamsuddin, S. M. ve Sudirman, R., EEG Signals Classification Using a Hybrid Method Based on Negative Selection and Particle Swarm Optimization, Machine Learning and Data Mining in Pattern Recognition, 7376 (2012) 427-438.
80. Tezel, G., Arslan, S. ve Isik, H, EEG Signals Classification Using a Hybrid Structure of ANN and PSO, International Journal of Future Computer and Communication, 1, 2 (2012) 170-172.
81. Kentucky, U. O., The Borda Count Method, The Mathematics of Voting, U.S.A, 2012.
82. Kutlu, F. ve Köse, C., Deneysel Epilepsi ile Elde Edilen ECoG İşaretlerinden Epileptik Nöbetin Algılanması, IEEE 22. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Nisan 2013, KKTC.

83. Kutlu, F. ve Köse, C., Epileptik Nöbetin ECoG İşaretlerinin Zaman Uzayı Özniteliklerinden Faydalanarak Algılanması, Çankaya Üniversitesi 6. Mühendislik ve Teknoloji Sempozyumu, Nisan 2013, Ankara, Bildiriler Kitabı: 187-190.

ÖZGEÇMİŞ

03.09.1988 tarihinde Samsun'un Havza ilçesinde dünyaya gelmiştir. İlkokul ve ortaokulu Vezirköprü Atatürk İlköğretim Okulu'nda ve liseyi Vezirköprü Anadolu Lisesi'nde tamamlamıştır. 2007 yılında Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü'nü kazanmış ve 2011 yılında mezun olmuştur. Aynı yıl güz döneminde yüksek lisans eğitimine başlamış, bahar döneminde de Amasya Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü'ne araştırma görevlisi olarak atanmıştır. 35. madde kapsamında halen Karadeniz Teknik Üniversitesi'nde araştırma görevlisi olarak çalışmaktadır. İyi derecede İngilizce bilmektedir.