

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**MAKİNE ÖĞRENME Sİ TEKNİKLERİ İLE ARİTMİ TESPİTİ VE YENİ
ÖZNİTELİKLER İLE BAŞARIMIN ARTIRILMASI**

DOKTORA TEZİ

Yasin KAYA

**KASIM 2017
TRABZON**



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE ARİTMİ TESPİTİ VE YENİ ÖZNİTELİKLER
İLE BAŞARIMIN ARTIRILMASI**

Yasin KAYA

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"DOKTOR (BİLGİSAYAR MÜHENDİSLİĞİ)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 29 / 09 / 2017

Tezin Savunma Tarihi : 16 / 11 / 2017

Tez Danışmanı : Yrd. Doç. Dr. Hüseyin PEHLİVAN

Trabzon 2017

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**Bilgisayar Mühendisliği Anabilim Dalında
Yasin KAYA Tarafından Hazırlanan**

**MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE ARİTMİ TESPİTİ VE YENİ
ÖZNİTELİKLER İLE BAŞARIMIN ARTIRILMASI**

**başlıklı bu çalışma, Enstitü Yönetim Kurulunun 10 /10/2017 gün ve 1722 sayılı
kararıyla oluşturulan jüri tarafından yapılan sınavda
DOKTORA TEZİ
olarak kabul edilmiştir.**

Jüri Üyeleri

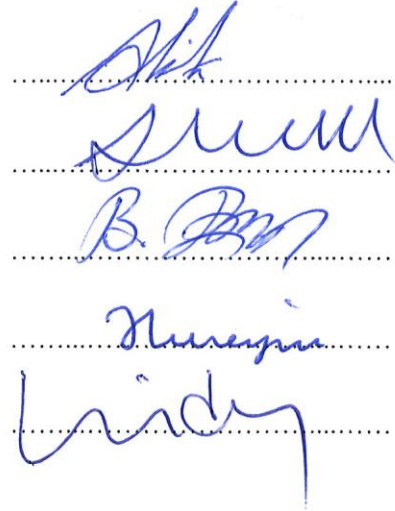
Başkan : Prof. Dr. Ali KARCI

Üye : Prof. Dr. Ali GANGAL

Üye : Doç. Dr. Burhan ERGEN

Üye : Yrd. Doç. Dr. Hüseyin PEHLİVAN

Üye : Yrd. Doç. Dr. Bekir DİZDAROĞLU



Prof. Dr. Sadettin KORKMAZ

Enstitü Müdürü

ÖNSÖZ

Kalp hastalıkları dünyadaki en önemli ölüm nedenlerinden biridir. Bu nedenle erken tanı ve teşhis için kalp vurularının izlenmesi, ilişkili tıbbi personel için hayati önem taşımaktadır. EKG vurularının farklı türlere göre otomatik alt kategorilere ayırarak ve bilgisayar destekli tanıya olanak sağlayarak, kardiyoloji doktorlarının bu kayıtların analizinde harcadıkları süreyi azaltmak mümkün olacaktır. Bu tez çalışmasında EKG aritmilerin sınıflandırılması için etkin yaklaşımlar geliştirilmiş ve önerilen yeni özniteliklerle aritmi sınıflandırma çalışması yapılmıştır.

Çalışmalarında danışmanlığımı üstlenen, her türlü destek ve katkılarıyla çalışmamı yönlendiren değerli danışman hocam Sayın Yrd. Doç. Dr. Hüseyin PEHLİVAN'a teşekkürlerimi sunarım.

Çalışma süresince değerli görüş ve katlılarını esirgemeyen sayın Prof. Dr. Ali GANGAL'a ve Yrd. Doç. Dr. Bekir DİZDAROĞLU'na teşekkür ederim. Çalışma fikir aşamasında iken destek ve görüşlerini esirgemeyen Tıbbi Bilişim Anabilim Dalı Başkanı sayın Prof. Dr. Kemal TURHAN'a teşekkür ederim.

Tez çalışması sürecinde bana her türlü desteklerini sunan Enformatik Bölümünden değerli mesai arkadaşlarım Öğr. Gör. Özkan BİNGÖL, Okt. Akif SÖNMEZ, Okt. Fatih GÜRCAN ve Öğr. Gör. Ercüment YILMAZ'a, İstatistik ve Bilgisayar Bilimleri Bölümünden arkadaşlarım Yrd. Doç. Dr. Uğur ŞEVİK ve Yrd. Doç. Dr. Tolga BERBER'e teşekkürü borç bilirim.

Öğrenim hayatım boyunca beni her zaman destekleyen, bugünlere getiren sevgili anneme ve babama şükranlarımı sunarım. Tez çalışmam süresince her zaman bana destek veren, bu çetrefilli süreçte tahammül eden sevgili eşim Hüsne KAYA'ya ve tezimin son aşamasında hayatımıza renk katan, verdiği pozitiflikle enerji kaynağımız biricik kızımız Ece Masal KAYA'ya sonsuz teşekkürlerimi sunarım.

Yasin KAYA

Trabzon 2017

TEZ ETİK BEYANNAMESİ

Doktora Tezi olarak sunduđum “Makine Öğrenmesi Teknikleri ile Aritmi Tespiti ve Yeni Öznitelikler ile Başarımın Artırılması” başlıklı bu çalışmayı baştan sona kadar danışmanım Yrd. Doç. Dr. Hüseyin PEHLİVAN'ın sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 16/11/2017

Yasin KAYA

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET	IX
SUMMARY	X
ŞEKİLLER DİZİNİ.....	XI
TABLolar DİZİNİ.....	XIV
SEMBOLLER DİZİNİ	XVI
1. GENEL BİLGİLER.....	1
1.1. Giriş.....	1
1.2. Çalışmanın Amacı	2
1.3. Tezin Kapsamı.....	4
1.4. Elektrokardiyogram.....	5
1.4.1. EKG'nin Kısa Tarihi	5
1.4.2. Kalbin Anatomik Yapısı.....	8
1.4.3. Kalbin İletim Sistemi.....	8
1.4.4. EKG İşaretinin Karakteristik Özellikleri.....	12
1.4.4.1. Taban Hattı Sapması	12
1.4.4.2. Elektromiyogram Gürültüsü.....	12
1.4.4.3. Güç Hattı Sapması.....	13
1.4.4.4. Elektrot Kontak Gürültüsü	13
1.4.4.5. Hareketten Kaynaklanan Gürültüler.....	13
1.4.5. EKG'ye Genel Bir Bakış	13
1.4.5.1. P Dalgası.....	15
1.4.5.2. PR Aralığı.....	15

1.4.5.3. QRS Kompleksi.....	16
1.4.5.4. ST Aralığı.....	16
1.4.5.5. QT Aralığı.....	16
1.4.5.6. RR Aralığı.....	17
1.4.6. EKG İşaretlerinin Kayıt Edilmesi.....	17
1.4.7. Klinik EKG Analizine Temel Bir Yaklaşım.....	19
1.4.8. Kardiyak Bozukluklara Genel Bir Bakış.....	20
1.5. EKG İşaretlerinin İşlenmesi.....	24
1.5.1. Domenler.....	24
1.5.1.1. Zaman Domeninde EKG İşaretleri.....	24
1.5.1.2. Frekans Domeninde EKG İşaretleri.....	25
1.5.1.3. Dalgacık Domeninde EKG İşaretleri.....	25
1.5.2. R Tepesinin Algılanmasında Kullanılan Algoritmalar.....	26
1.5.2.1. Türev ve Dijital Filtre Tabanlı Algoritmalar.....	27
1.5.2.2. Dalgacık Tabanlı Algoritmalar.....	28
1.6. EKG İşaretlerinde Öznitelik Çıkarımı.....	28
1.6.1. Zaman Domeninde Öznitelik Çıkarımı ve Analizi.....	29
1.6.2. Frekans Domeninde Öznitelik Çıkarımı ve Analizi.....	30
1.6.3. Dalgacık Domeninde Öznitelik Çıkarımı ve Analizi.....	31
1.7. Boyut Azaltma.....	32
1.7.1. Öznitelik Seçimi.....	33
1.7.1.1. Sıralı Arama Algoritmaları ile Öznitelik Seçimi.....	35
1.7.1.2. Genetik Algoritma ile Öznitelik Seçimi.....	36
1.7.2. Öznitelik Azaltma.....	39
1.7.2.1. Temel Bileşen Analizi.....	40
1.7.2.2. Bağımsız Bileşen Analizi.....	42
1.7.2.3. Özdüzenleyici Haritalar.....	45
1.8. EKG Verilerinin Sınıflandırılması.....	49
1.8.1. K-En Yakın Komşu Algoritması.....	50

1.8.2.	Yapay Sinir Ağları.....	52
1.8.2.1.	Yapay Sinir Ağlarında Öğrenme.....	54
1.8.3.	Destek Vektör Makinesi.....	56
1.8.3.1.	Verilerin Doğrusal Olarak Ayrılama Durumu.....	61
1.8.3.2.	DVM ile Çok Sınıflı Sınıflandırma.....	66
1.8.4.	Karar Ağaçları ile Sınıflandırma.....	66
1.9.	Sınıflandırma Başarımının Değerlendirilmesi İçin Performans Ölçütleri.....	68
1.10.	Sınıflandırma İçin Veri Setinin Organizasyonu.....	72
1.11.	Literatür Çalışması.....	73
1.11.1.	Erken Karıncık Kasılması Virüsünün Sınıflandırılması ile İlgili Literatür Değerlendirmesi.....	73
1.11.2.	EKG Vurularının Sınıflandırılması ile İlgili Literatür Değerlendirmesi.....	75
2.	YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME.....	78
2.1.	MIT-BIH Aritmi Veritabanı.....	79
2.2.	Ön İşleme.....	81
2.3.	QRS Tespiti.....	84
2.4.	Vuru Ayırıştırma.....	90
2.5.	Erken Karıncık Kasılması Virüsünün Algılanması.....	91
2.5.1.	Erken Karıncık Kasılması Virüsünün Algılanması İçin Veri Seti.....	92
2.5.2.	EKK Sınıflandırma İçin Boyut İndirgeme.....	94
2.5.3.	Genetik Algoritma ile Öznitelik Seçimi.....	96
2.5.4.	EKK Vurularının Sınıflandırılması.....	98
2.5.5.	EKK Vurularının Sınıflandırma Sonuçları.....	99
2.5.6.	TB ve BB Sayısının Sınıflandırma Başarımına Etkisi.....	102
2.5.7.	Sınıflandırma Hesaplama Süreleri.....	103
2.6.	EKG Vurularının Sınıflandırılması.....	104
2.6.1.	Testlerde Kullanılan Veri Seti.....	105
2.6.2.	Öznitelik Hesaplama.....	108
2.6.2.1.	Çarpıklık ve Basıklık.....	109
2.6.2.2.	Diğer İstatistiksel Özniteliklerin Hesaplanması.....	111

2.6.2.3. RR Aralıklarının Hesaplanması.....	113
2.6.3. EKG Vurularının Sınıflandırılması İçin Öznitelik Boyut İndirgeme	115
2.6.3.1. TBA, BBA ve ÖH Kullanılarak Boyut İndirgeme	116
2.6.3.2. GA, SİA ve SGA Yöntemleri ile Öznitelik Seçme	122
2.6.4. Sınıflandırma	126
2.6.4.1. YSA ile Sınıflandırma	129
2.6.4.2. K-NN ile Sınıflandırma	130
2.6.4.3. DVM ile Sınıflandırma.....	131
2.6.4.4. KA ile Sınıflandırma	132
2.6.5. Sınıflandırma Sonuçları.....	134
2.6.6. Büyük Veri Testi Sonuçları	139
2.6.7. Öznitelik Azaltma ve Sınıflandırma Hesaplama Süreleri	140
3. SONUÇLAR	142
3.1. EKK Vurusunun Sınıflandırılması ile İlgili Sonuçlar ve Değerlendirme	142
3.2. Dokuz Ayrı EKG Vurusunun Sınıflandırılması ile İlgili Sonuçlar ve Değerlendirme	145
4. ÖNERİLER	151
5. KAYNAKÇA	152
ÖZGEÇMİŞ	

Doktora Tezi

ÖZET

MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE ARİTMİ TESPİTİ VE YENİ
ÖZİNİTELİKLER İLE BAŞARIMIN ARTIRILMASI

Yasin KAYA

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Yrd. Doç. Dr. Hüseyin PEHLİVAN
2017, 167 Sayfa

Kardiyak aritmiler kalp hastalıklarının en önemli göstergesinden bir tanesidir. Aritmiler kalbin, iletim, atım oranı veya elektriksel aktivitesinde ortaya çıkabilecek aksamalardan kaynaklanmaktadır. Erken karıncık kasılması (EKK) ise ektopik kalp vuruları ile ortaya çıkan en yaygın kardiyak aritmilerinden biridir. Aritmilerin algılanması kalple ilgili olabilecek problemlerin önceden tahmininde önemli rol oynamaktadır. Bu çalışmada ilk olarak işarete ait zaman serilerini kullanarak EKK vurusunun sınıflandırılması gerçekleştirilmiştir. Bir vuruluk işaretin zaman serisine ek olarak çeşitli boyut indirgeme algoritmalarının performans etkisi araştırılmıştır. Ek olarak çalışma daha geniş kapsamda aritmi türlerini sınıflandıracak şekilde genişletilmiş ve testler gerçekleştirilmiştir. Bu aritmilerin sınıflandırılması için bir vuruluk işarete ait yeni öznelikler çıkarılmıştır. Öznelikler, boyut indirgeme algoritmaları kullanılarak daha küçük boyutlara düşürülmüştür. Deneyler, k-en yakın komşu algoritması, yapay sinir ağları, karar ağaçları ve destek vektör makinesi sınıflandırıcıları kullanılarak gerçekleştirilmiştir. Bulgular, doğruluk, duyarlılık, kesinlik, özgünlük ve çalışma süreleri bakımından değerlendirilmiştir. Çalışmada yapılan testlerde kullanılan veriler bu alanda standart hale gelmiş olan MIT-BIH aritmi veritabanından alınmıştır.

Anahtar Kelimeler: EKG işaret analizi, Erken Karıncık Kasılması, EKG aritmi sınıflandırma, Öznelik Boyut indirgeme, İşaret işleme, Makine Öğrenmesi

PhD. Thesis

SUMMARY

ARRHYTHMIA DETECTION WITH MACHINE LEARNING TECHNIQUES AND
INCREASING THE CLASSIFICATION PERFORMANCE USING NEW FEATURES

Yasin KAYA

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Computer Engineering Graduate Program
Supervisor: Asst. Assoc. Dr. Hüseyin PEHLİVAN
2017, 167 Pages

Cardiac arrhythmia is one of the most important indicators of heart disease. Heart arrhythmias are caused by any disruption in the regularity, rate, or transmission of the cardiac electrical impulse. Premature ventricular contractions (PVCs) are a common form of cardiac arrhythmia caused by ectopic heartbeats. The detection of arrhythmias by means of ECG (electrocardiogram) signals is important for the prediction of possible heart failure. This study was first focused on PVC classification using time series of ECG signal. Moreover, the performance effects of several dimension reduction approaches were also tested. In addition, the work is extended to classify more common arrhythmia types. Statistical features are calculated from one beat signal for the classification. Feature size were reduced to a lower size using size reduction algorithms. Experiments were carried out using well-known machine learning methods, including neural networks, k-nearest neighbor, decision trees, and support vector machines. Findings were expressed in terms of accuracy, sensitivity, specificity, selectivity and running time. The data used in the tests in this study were taken from the MIT-BIH arrhythmia database that has become the standard in this area.

Key Words: ECG signal analysis, Premature Ventricular Contraction, ECG arrhythmia classification, Feature size reduction, Signal processing, Machine learning.

ŞEKİLLER DİZİNİ

Sayfa No

Şekil 1.	Elektrokardiyograf [8].	6
Şekil 2.	Einthoven Modeli	7
Şekil 3.	Kalbin anatomik yapısı [1].	8
Şekil 4.	Kalbin iletim sistemi [1].	9
Şekil 5.	EKG dalga şekli ve üzerindeki önemli dalgalar [1].	11
Şekil 6.	Kardiyak döngüsü ve EKG ilişkisi [1].	11
Şekil 7.	EKG'deki şekilsel öznitelikler	14
Şekil 8.	Derivasyonlar, (a) Kalp bölgesine ait elektrotların yerleşimi, (b) Standart derivasyonlar.	17
Şekil 9.	Dalgacık ayrıştırma ağacı. Y: Sinyalin yaklaşım bilgisi D: Sinyalin detay bilgisi	31
Şekil 10.	Öznitelik Seçimi	34
Şekil 11.	Filtre yöntemleri ile öznitelik seçimi	34
Şekil 12.	Paketleyici yöntemler ile öznitelik seçimi	35
Şekil 13.	GA genel yapısı	37
Şekil 14.	Kromozomların gösterimi. (a) ikili kodlama, (b) gerçel kodlama	37
Şekil 15.	İkili diziler için standart tek nokta çaprazlama	38
Şekil 16.	İkili dizi için iki bit mutasyon	38
Şekil 17.	GA ile öznitelik seçimi	39
Şekil 18.	Öznitelik çıkarımı	39
Şekil 19.	Örnek veri ve hesaplanmış 1. Temel bileşenin yönü	41
Şekil 20.	ÖH Genel yapısı	46
Şekil 21.	Kazanan birimin komşulukları	47
Şekil 22.	Komşuluk çemberinin daraltılması	48
Şekil 23.	K-NN sınıflandırma örneği	51
Şekil 24.	Yapay sinir hücresinin yapısı	53

Şekil 25.	Çok katmanlı ağ.....	55
Şekil 26.	İki boyutlu uzayda doğrusal olarak ayrılabilen veriler	57
Şekil 27.	Doğrusal olarak birbirinden ayrılabilen eğitim verileri için muhtemel ayırıcı düzlem.....	58
Şekil 28.	ξ gevşek değişkenlerinin ifadesi	62
Şekil 29.	Doğrusal olarak ayıramayan veri örnekleri	63
Şekil 30.	X ve Y değişkenleri için basit bir karar ağacı.....	67
Şekil 31.	Veri kümesinin organizasyonu	72
Şekil 32.	Doğruluğun tahmin edilmesi	72
Şekil 33.	k=4 için veri seti	73
Şekil 34.	Vuru sınıflandırma sistemin genel blok diyagramı	79
Şekil 35.	100 numaralı kaydın ilk 3 saniyelik görünümü	80
Şekil 36.	Medyan Filtre. Üst: pencere genişliği 3 olan filtre sonucu, Alt: pencere genişliği 10 olan filtre sonucu.....	82
Şekil 37.	Kademeli Alçak geçiren filtre uygulamasının blok diyagramı.....	83
Şekil 38.	Taban hattı sapmasının filtrelenmesi.	84
Şekil 39.	QRS Algılama algoritmasının genel şeması.....	85
Şekil 40.	Orijinal EKG ve 5-15 Hz bandında filtrelenmiş sonucu	86
Şekil 41.	Ham EKG ve Filtrelenen sinyalin genlik spektrumu.....	86
Şekil 42.	Türev, Kare İşlemi ve Kayan Pencere işlemleri	87
Şekil 43.	QRS algılama işlemi sonucu.....	89
Şekil 44.	100 numaralı kaydın ilk 10 vurusu	90
Şekil 45.	EKK vurusunun sınıflandırılmasını gösteren şema	92
Şekil 46.	Temel bileşenlerin birikimli varyans değerleri.....	95
Şekil 47.	ÖH ile boyutu indirgenmiş test verisi.....	96
Şekil 48.	Bir vuruluk sinyalin zaman serisi ve GA ile seçilmiş örnekleri	97
Şekil 49.	K sayısının sınıflandırma başarımına etkisi.....	102
Şekil 50.	TB ve BB sayısının sınıflandırma başarımına etkisi	103
Şekil 51.	EKG vuru sınıflandırma deneyinin genel blok diyagramı.....	105
Şekil 52.	Veri setindeki her bir vuru tipine ait ilk 50 örnek	107

Şekil 53.	Vuruya ait özniteliklerin hesaplanması	109
Şekil 54.	Negatif ve Pozitif Çarpıklık.....	110
Şekil 55.	Basıklık gösterimi.....	111
Şekil 56.	Çeyrekler aralığının hesaplanması.....	112
Şekil 57.	R-R Aralıklarının hesaplanması	113
Şekil 58.	Bir vuruluk sinyal için Çarpıklık ve Basıklık öznitelikleri.....	114
Şekil 59.	Önceki R-R aralığı ve Sonraki R-R Aralığı öznitelikleri	115
Şekil 60.	TB'in birikimli varyans değerleri	116
Şekil 61.	İlk iki temel bileşenin test verisinin temsili.....	118
Şekil 62.	İlk iki bağımsız bileşenin test verisini temsili	119
Şekil 63.	ÖH kullanılarak öznitelik uzayının iki boyuta indirgenmiş hali	121
Şekil 64.	ÖH kullanılarak öznitelik uzayının 3 boyutlu uzaya indirgenmiş hali.....	122
Şekil 65.	EKG vurularını sınıflandıran örnek Karar Ağacı	134
Şekil 66.	K sayısının sınıflandırma başarımına etkisi.....	137

TABLolar DİZİNİ

Sayfa No

Tablo 1.	Dakikada 60 vuru atan sağlıklı erkeğin II derivasyon sinüs ritmine ait EKG özellikleri [13].	15
Tablo 2.	AAMI standartlarına göre vuru tipleri	21
Tablo 3.	Performans Ölçütleri	69
Tablo 4.	Karşıtlık Matrisi	69
Tablo 5.	EKK sınıflandırma için veri setindeki vuruların dağılımı	93
Tablo 6.	YSA gizli katman düğüm sayısı ve farklı giriş vektörlerinin başarıma etkisi..	98
Tablo 7.	Giriş verisi olarak zaman serisi için sınıflandırma sonuçları (%)	100
Tablo 8.	Giriş verisi olarak temel bileşenler için sınıflandırma sonuçları (%)	100
Tablo 9.	Giriş verisi olarak bağımsız bileşenler için sınıflandırma sonuçları (%)	100
Tablo 10.	Giriş verisi olarak ÖH tarafından üretilen öznitelikler için sınıflandırma sonuçları (%)	100
Tablo 11.	Giriş verisi olarak GA tarafından seçilen öznitelikler için sınıflandırma sonuçları (%)	101
Tablo 12.	Sınıflandırıcıların kullandıkları giriş vektörüne göre sınıflandırma süreleri (sn)	104
Tablo 13.	Testlerde kullanılan vuru türleri ve örnek sayıları	106
Tablo 14.	Hesaplanan Öznitelikler	114
Tablo 15.	TB sayısı ve sınıflandırma başarıımı	117
Tablo 16.	BB Sayısı ve sınıflandırma başarıımı	120
Tablo 17.	GA tarafından seçilen öznitelikler	123
Tablo 18.	SİA algoritması tarafından seçilen öznitelikler	125
Tablo 19.	SGA tarafından seçilen öznitelikler	126
Tablo 20.	Çok sınıflı sınıflandırma için karşıtlık matrisi	127
Tablo 21.	Öznitelik grupları için YSA gizli katman düğüm sayılarının sınıflandırma başarıma etkisi	130
Tablo 22.	Öznitelik grupları için K-NN algoritmasında K değerinin başarıma etkisi	131

Tablo 23.	Lineer çekirdek fonksiyonu için C parametresinin başarıma etkisi.....	132
Tablo 24.	Seçilen çekirdek fonksiyonlarının başarıma etkisi	132
Tablo 25.	KA sınıflandırmada parametre optimizasyonu.....	133
Tablo 26.	Giriş verisi olarak 32 istatistiksel öznitelik için sınıflandırma sonuçları (%)	135
Tablo 27.	Giriş verisi olarak TB'ler için sınıflandırma sonuçları (%).....	135
Tablo 28.	Giriş verisi olarak BB'ler için sınıflandırma sonuçları (%).....	135
Tablo 29.	Giriş verisi olarak ÖH için sınıflandırma sonuçları (%).....	136
Tablo 30.	Giriş verisi olarak GA ile seçilen öznitelikler için sınıflandırma sonuçları (%)	136
Tablo 31.	Giriş verisi olarak SİA ile seçilen öznitelikler için sınıflandırma sonuçları (%)	136
Tablo 32.	Giriş verisi olarak SGA ile seçilen öznitelikler için sınıflandırma sonuçları (%)	137
Tablo 33.	10-kez çapraz doğrulama için her adımda elde edilen sınıflandırma başarıma	138
Tablo 34.	Büyük veri testi için vuruların sınıflara dağılımı ve oranları	139
Tablo 35.	Büyük veri sınıflandırma karşıtlık matrisi.....	140
Tablo 36.	Farklı giriş vektörleri için sınıflandırma algoritmalarının hesaplama süreleri (sn)	140
Tablo 37.	Öznitelik azaltma algoritmalarının hesaplama süreleri (sn.)	141
Tablo 38.	Çalışmada kullanılan sınıflandırıcılar ve yayınlanmış EKK sınıflandırıcılarının yazarları tarafından raporlanan performans metrikleri (%)	143
Tablo 39.	Çalışmada önerilen yöntem ve yayınlanmış EKG aritmi sınıflandırıcılarının yazarları tarafından raporlanan performans metrikleri (%)	149

SEMBOLLER DİZİNİ

a	:Anormal Kulakçık Erken Vurusu
A	:Dağılım Aralığı
AHA	:Amerika Kalp Derneği
AV	:Atriyo-Ventriküler
aVF	:Unipolar Sol Bacak Derivasyonu
aVL	:Unipolar Sol Kol Derivasyonu
aVR	:Unipolar Sağ Kol Derivasyonu
B	:Basıklık
BB	:Bağımsız Bileşenler
BBA	:Bağımsız Bileşen Analizi
BDT	:Bilgisayar Destekli Tanı
BMU	:Best Matching Unit
C	:Çarpıklık
Ç	:Çeyrekler Aralığı
ÇDD	:Çapraz Doğrulama Doğruluğu
ÇDOD	:Çapraz Doğrulama Ortalama Duyarlılığı
ÇDOK	:Çapraz Doğrulama Ortalama Kesinliği
ÇKA	:Çok Katmanlı Ağ
DDA	:Doğrusal Diskriminant Analizi
DSÖ	:Dünya Sağlık Örgütü
DVM	:Destek Vektör Makinesi
E	:Karıncık Kaçak Vurusu
$E\{.\}$:İstatistiksel Beklenen Değer
EKG	:Elektrokardiyogram
EKK	:Erken Karıncık Kasılması
EMG	:Elektromiyogram
f	:Yapay ve Normal Vuru Birleşimi
F	:Karıncık ve Normal Vuru Birleşimi
FN	:Yanlış Negatif

FP	:Yanlış Pozitif
GA	:Genetik Algoritmalar
GUR_ESK	:Gürültü Eşik Deęeri
GUR_SEV	:Gürültü Seviye Deęeri
HFD	:Hızlı Fourier Dönüşümü
KA	:Karar Ağaçları
KAH	:Koroner Arter Hastalığı
KB	:Kazanan Birim
KDH	:Kalp ve damar hastalığı
KH	:Kalp Hızı
KHD	:Kalp Hızı Deęişkenliği
K-NN	:K-En Yakın Komşu Algoritması
LBBS	:Left Bundle Branch Block
L	:Sol Dal Blok Vurusu
MD	:Maksimum Derinlik
med{.}	:Medyan Filtre
MYS	:Minimum Yaprak Sayısı
N	:Normal Vuru
OD	:Ortalama Duyarlılık
OK	:Ortalama Kesinlik
OKH	:Ortalama Karesele Hata
ÖH	:Özdüzenleyici Haritalar
P, Q, R, S, T	:EKG Üzerindeki Genlik Sapmaları
R	:Saę Dal Blok Vurusu
PR	:EKG'deki PR Aralığı
QRS	:EKG üzerindeki QRS kompleksi
RBBS	:Right Bundle Branch Block
RBF	:Radial Basis Function
RR	:EKG'deki RR Aralığı
SA	:Sino-Atriyal
SGA	:Sıralı Geri Arama
SIN_ESK	:İşaret Eşik Deęeri

SIN_SEV	:İşaret Seviye Değeri
SİA	:Sıralı İleri Arama
$S_{n \times n}$: $n \times n$ Kovaryans Matrisi
ST	:EKG'deki ST Aralığı
TB	:Temel Bileşenler
TBA	:Temel Bileşen Analizi
TDA	:Tekil Değer Ayrışımı
TEPEI	:EKG İşaretindeki Tepe Noktası
TN	:Doğru Negatif
TP	:Doğru Pozitif
u	:Destek vektör ve ayırıcı düzlem arasındaki uzaklık
V	:Erken Karıncık Kasılması Vurusu
W	:Ağırlık Vektörü
x(t), y(t)	:EKG İşaretin Zaman Serisi
Y(t)	:Medyan Filtre ile Filtrelenmiş İşaret
YSA	:Yapay Sinir Ağları
/	:Yapay Vuru
\bar{x}	:Vektörel Ortalama
λ_i	:Kovaryans Matrisinin Özdeğerleri
$\sigma(t)$:t Anındaki Komşuluk Yarıçapı
$L(t)$:t Anındaki Öğrenme Oranı
$\theta(t)$:t Anındaki Düğümün Kazanan Düğümünden Olan Uzaklığının Öğrenmeye Olan Etkisi
E_{MSE}	:Ortalama Kare Hata
$Q(w, b, \alpha)$:İkili Optimizasyon Problemi
ξ_i	:Ayırıcı Düzlem Hatası Gevşek değişkeni
$K(x_i, x_j)$:Çekirdek Fonksiyonu

1. GENEL BİLGİLER

1.1. Giriş

Ortalama bir yetişkin bireyin kalp atım oranı dakikada ortalama 75 atım olarak kabul edilirse, bir insan kalbi 75 yıllık ömrü boyunca günde yaklaşık 108.000 kez, bir yılda 39 milyondan fazla ve toplamda yaklaşık olarak 3 milyar kez kasılmış olacaktır. Dinlenme durumunda bir yetişkinin kalbi her kasılma başına yaklaşık 70ml kan pompalar. Bu sıvı dakika başına 5,25 litre ve günde yaklaşık 7500 litreye denk gelecektir. Bir yıl boyunca kalbin pompaladığı kan miktarı ise yaklaşık 3 milyon litre olacaktır [1]. Bu şekilde çalışan bir organın çalışma şeklinin incelenmesi ve olabilecek problemlerin önceden tespit edilerek tedavisinin uygulanması hayati önem taşımaktadır.

Geçtiğimiz yüzyılda dünyada eğitim ve gelir düzeyindeki yükselme, beslenme alışkanlıklarının değişmesi, bulaşıcı hastalıkların kontrolü gibi etkenler beklenen yaşam süresinin uzamasına neden olmuştur. Yaşam süresinin uzaması paralelinde bulaşıcı olmayan kronik hastalıkların görülme sıklığında artış meydana gelmiştir. Yaşlı nüfusun çocuk nüfusa oranla artması, toplumdaki sağlık sorunlarının çocukluk çağı hastalıklarından yaşlı nüfusta görülen bulaşıcı olmayan kronik hastalıklarına yaygınlaşmasına yol açmıştır [2]. Kalp ve damar hastalıkları, küresel ölçekte en önemli ölüm sebeplerinden biridir. Dünya Sağlık Örgütü (DSÖ) verilerine göre 2005 yılında 17,5 milyon insanın kalp ve damar hastalıkları sebebiyle öldüğü ve bu rakamın küresel ölümlerin %30'unu oluşturduğu tahmin edilmektedir. Bu ölümlerin 7,6 milyonu kalp krizlerine bağlıdır. Ölümlerin %80'i düşük ve orta gelirli ülkelerde meydana gelmiştir. Eğer uygun önlemler alınmazsa 2020 yılına kadar her yıl yaklaşık 20 milyon insan daha kalp ve damar hastalıklarından, özellikle de kalp krizleri ve diğer kalp rahatsızlıkları sebebiyle hayatını kaybedecektir [2]. Amerika Kalp Derneği verileri (AHA), 83,6 milyon Amerikalı yetişkinin bir veya daha fazla kalp ve damar hastalığına sahip olduğunu ortaya koymaktadır. Bunlardan 42.2 milyonunun 60 yaşın üstünde olduğu tahmin edilmektedir. Kalp ve damar hastalıkları (KDH) 2010 yılında Amerika'da gerçekleşen ölümlerin %31,9'unu oluşturarak ölüm nedenlerinin başında gelmektedir [3].

Yukarıda verilen istatistiklerden görüleceği gibi, kalp hastalıkları dünyadaki ölüm nedenlerinde başı çekmektedir. Bu nedenle erken tanı ve teşhis için kalp döngülerinin izlenmesi ve analizi kardiyoloji doktorları ve ilgili tıbbi personel için hayati önem taşımaktadır. Kalple ilgili hastalıkların teşhisinde elektrokardiyogram (EKG) kayıtlarının hızlı bir şekilde yorumlanması önem arz etmektedir. Kalbin fizyolojik ve çalışma durumunda oluşabilecek hatalar hakkında yeterli bilginin elde edilmesi uzun kayıtların analizi sonucunda ortaya çıkmaktadır. EKG vurularının bazı hastalıklar hakkında içerdiği bilgiye göre otomatik alt kategorilere ayırarak ve bilgisayar destekli tanıya (BDT) olanak sağlayarak, kardiyoloji doktorlarının ve ilgili personelin bu kayıtların analizinde harcadıkları süreyi azaltmak mümkün olacaktır. BDT’da alt kategorilere ayırma işlemi gerçek zamanlı olarak yapılacaksa geliştirilecek yöntemin işlem yükünün az olması gerekir. Ancak tüm EKG kayıtları ham olarak kaydedilip analizi bilgisayar ortamında yapılacaksa uygulanacak sınıflandırmanın mümkün olduğunca daha fazla sınıfa ayrılması BDT açısından işlem zamanından çok daha önemlidir.

Bilişim teknolojisindeki gelişim biyolojik işaretlerin işlenmesinde önemli rol oynamaktadır. Son yıllarda yüksek performanslı hesaplama tekniklerindeki büyük gelişim, biyolojik sistemlerde temel sorunlara çözüm üreten modellerin üretilmesinde yeni yaklaşımlar geliştirilerek önemli katkılarda bulunmuştur.

Elektrokardiyografi biyolojik işaretler arasında en çok çalışılan işaret tipidir. Onlarca yıldır EKG işaret araştırmaları bu temel disiplini kalp rahatsızlıklarının tanısı için bir araç haline getirmiştir. Basitliği, düşük maliyeti ve insan vücudu üzerinden hastaya zarar vermeden elde edilmesi nedeniyle yeni teknolojilere rağmen yaygın olarak kullanılmaktadır.

1.2. Çalışmanın Amacı

EKG kalbin çalışması sonucu ortaya çıkan, kalbin elektriksel aktivitesi hakkında bilgi veren ve tanı aracı olarak kullanılan en yaygın araçlardan biridir. EKG, hastaya zarar vermeden, insan vücudunun çeşitli noktalarına yerleştirilen elektrotlar ile kolayca elde edilebilir. Bu şekilde EKG ile kalbin çalışması sırasında ortaya çıkacak anormallikler kayıt altına alınmış olur. Bu kayıtlar işleme ve yorumlanma açısından büyük önem taşımakla birlikte, kardiyoloji doktorlarına büyük iş yükü getirmektedir. Tanı ve teşhis için

kullanılacak EKG kayıtları genellikle Holter kayıtlarıdır ve 24-48 saatlik zaman dilimlerinden oluşmaktadır. Kalp hakkında önemli bilgilerin elde edilebilmesi bu uzun kayıtların incelenmesiyle gerçekleşir. Bu işlemde genellikle tüm kayıtlar ham hali ile kaydedildikten sonra çevrim dışı olarak analiz edilmesi için saklanır ve geliştirilecek yöntem ile kalp vurularının analizi yapılır. Bazı durumlarda kaydı yapacak cihaza bütünlük bir yöntem cihazın kayda başlamasını tetikleyen bir vurunun tanınması ile sağlanabilir. Böylece yapılacak kaydın süresi azaltılmış olur.

Bu çalışma Holter kayıtlarından elde edilmiş uzun süreli EKG işaretlerinin analizini gerçekleştirmek için sınıflandırma algoritmaları ve öznelik azaltma yöntemlerini içeren alt yapı çalışmalarını kapsamaktadır. Belirli zamanlarda ortaya çıkan özel kalp olaylarını tanımlayan kardiyak aritmi analizi yöntemlerinin geliştirilmesi bu çalışmanın seçilmesindeki ana nedenlerden biridir. Bu olaylar hasta için zararlı olabilecek kardiyak hastalıklar ile ilişkilidir. Geliştirilecek yöntemlerin üretilen yeni cihazlar için bir alt yapı oluşturacağı göz önünde bulundurulduğunda, gerçek zamanlı tanıya olanak tanınması, kalp problemlerinden dolayı ortaya çıkacak ölüm oranlarının düşürülmesinde önemli rol oynayacaktır.

Tipik bir 24 saatlik EKG kaydı morfolojik olarak daha küçük gruplarda sınıflara ayrılması gereken yaklaşık 100.000 kalp vurusundan oluşmaktadır. Bu sınıfların çoğunda kalp atışının belirli bir deseni vardır. Kaç tane sınıf olduğunu ve sınıflandırma için kalp vurusunu temsil eden morfolojik şablonu bilmek yeterlidir. Kardiyak aktivite anomalisi veya hastalık belirtilerinin hangi zaman diliminde gerçekleştiğini bilmek için işaretin kaydının tamamını incelemek gerekmektedir.

Değerlendirme, kayıtlarda saklanan büyük miktarda verinin hangi tür kalp vurusuna ait olduğu ve belirli bir analiz için önemli olduğu, kararını verecek uzman için zor bir görev haline gelmektedir. Bazı patoloji veya ölümcül hastalıkların önlenmesi için tanıda önemli hale gelen tüm kaydın sadece birkaç vurusunun önemli olduğu durumlar vardır. Bu nedenle tüm kayıtların ayrıntılı olarak analizi gereklidir.

Bu çalışmada kardiyak aritmileri tespit etmek amacıyla ön işleme, öznelik çıkarımı, boyut azaltma ve sınıflandırma aşamaları da dâhil olmak üzere Holter kayıtlarının EKG işaretlerini analiz edebilmek için danışmanlı öğrenme tekniklerinin uygulanması

amaçlanmıştır. Tasarlanan yöntem ile MIT-BIH aritmi veritabanındaki vuru tiplerinden en yaygın olanların sınıflandırılması amaçlanmıştır.

Çalışmada iki ana konu üzerinde yoğunlaşmıştır. İlkinde erken karıncık kasılması (EKK) vurusun sınıflandırılması için bir model geliştirilmiştir. İkincisinde ise daha geniş kapsamda EKG vurularının sınıflandırılması için düşük hesaplama yükü ve yüksek başarımlı sağlayacak özniteliklerin hesaplanabileceği sınıflandırma modelleri oluşturulmaya çalışılmıştır. EKG vuruları için 9 farklı sınıf belirlenmiş ve bunları en iyi şekilde sınıflandırabilecek yeni özniteliklerin ve sınıflandırma modellerinin performansları analiz edilmiştir.

1.3. Tezin Kapsamı

Tüm örüntü tanıma yöntemlerinde olduğu gibi EKG vurularının sınıflandırılması da üç aşamalı bir işlemi gerektirir. İlk aşama; işaret üzerinde taşınan bilgiyi ortaya çıkarabilmek için çevresel faktörlerden arındırılacağı ön işleme aşamasıdır. İkinci aşama; işareti temsil edecek ve işaretin zaman serisinden daha küçük boyuta sahip öznitelik değerlerinin hesaplanması aşamasıdır. Son aşama ise geliştirilecek yöntem ve parametreler ile sınıflandırma işleminin gerçekleştirilmesi aşamasıdır [4].

Çalışmanın ilk bölümünde EKK vurusunun sınıflandırılması için işaretin zaman serisi sınıflandırma algoritmalarında giriş verisi olarak kullanılmıştır. Böylece sınıflandırmada harcanan süre azaltılmaya çalışılmıştır. İşlem süresini daha da hızlandırmak için bağımsız bileşen analizi, temel bileşen analizi gibi yöntemler kullanılarak öznitelik azaltma yoluna gidilmiş ve başarılı sonuçlar elde edilmiştir.

Çalışmanın ikinci bölümünde ise daha fazla sınıfta sınıflandırma yapabilmek için basit istatistik özelliklerden oluşan öznitelik vektörleri hesaplanmış, yine öznitelik azaltma yöntemleri ile öznitelik vektörlerinin boyutu düşürülmüş ve sınıflandırma algoritmaları kullanılarak kalp vuruları sınıflandırılmıştır.

Sınıflandırıcı olarak, yapay sinir ağları, k-en yakın komşu algoritması, destek vektör makinesi ve karar ağaçları sınıflandırma algoritması kullanılmış ve bunlardan elde edilen

sonular karşılařtırılmıřtır. Bu modellerin en iyi sonucu vermesi iin yapay zekâ yöntemleri kullanılarak parametre optimizasyonları yapılmıřtır.

1.4. Elektrokardiyogram

Kalbin alıřması sırasında oluřan iřaretin vücut üzerine yerleřtirilen elektrotlar vasıtasıyla kaydedilmesi iřlemine Elektrokardiyografi (EKG), kaydı gerekleřtiren cihaza ise Elektrokardiyograf denir. EKG tanıda kullanılan en iyi bilinen biyolojik iřaretlerden biridir. EKG kalp ile ilgili anormalliklerin tanısında, oluřabilecek fizyolojik problemlerin önceden teřhisinde ve kalpte meydana gelen deęiřikliklerin kayıt altına alınmasında kullanılır [5]. EKG iřareti vurularla tekrar etse de kaydedilen EKG'deki kalp vuru oranı zamanla deęiřir. Vuru oranının ortalaması ve varyansı zamanda deęiřiklik gösterir. Bu yüzden EKG iřareti yarı periyodik ve duraęan olmayan (non-stationary) olarak kabul edilir [6].

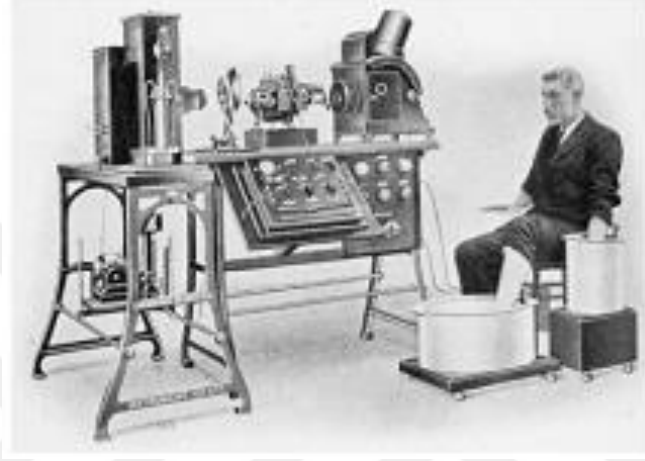
1.4.1. EKG'nin Kısa Tarihi

İlk insan EKG'si Thompson sifon kaydedici kullanarak 1869 yılında Alexander Muirhead tarafından kaydedilmiřtir. İnsan EKG'si ilk olarak 1897 yılında St. Mary Tıp Fakültesinde İngiliz fiziki Augustus D. Waller tarafından tanımlanmıřtır. Bu alıřmadan önce bir elektrometre yalnızca açık bir hayvan kalbi üzerinde kullanılmıřtı. Waller insan uzuvları üzerinden EKG kayıtlarının yapılabileceęini fark eden ilk kiři olmuřtur. Aynı zamanda "Elektrokardiyogram" kelimesini kullanan ilk kiřidir [6].

1895 yılında Hollandalı fizyolog Willem Einthoven ilk EKG makinesini geliřtirmiřtir [7]. Daha sonra 1901 yılında kalbin elektriksel aktivitesini güvenilir ve doęrudan bir yöntem ile kaydedilmesini saęlayan dizge galvanometreyi (string galvanometer) keřfetmiřtir [6]. Geliřtirilen elektrokardiyograf Őekil 1'de gösterilmiřtir. Őekilde elektrotlar hastanın kollarına ve ierisinde tuzlu su özeltilisi bulunan kovaya daldırılmıř bir ayaęına yerleřtirilmiřtir.

Einthoven günümüzde hala kullanılan EKG'deki sapmaları P, Q, R, S ve T olarak adlandırmıřtır. P dalgası, QRS kompleksi ve T dalgası sırasıyla kulakık kasılmasını,

karıncık kasılmasını ve karıncık gevşemesini göstermektedir. U dalgası ise Einthoven tarafından daha sonra galvanometre ile tanımlanmıştır. U dalgasının sebebi hala net olarak açıklanamamış değildir. Şu eylemlerin U dalgasının sebebi olduğu düşünülmektedir: Purkinje liflerinin repolarizasyonu, orta miyokardiyal hücrelerin (M cells) uzun aksiyon potansiyelleri veya karıncık alanındaki geç gevşemeden kaynaklanan gecikmiş repolarizasyon.



Şekil 1. Elektrokardiyograf [8].

Günümüzde Einthoven Üçgeni olarak bilinen standart derivasyon I, II ve III'ü gösteren eşkenar üçgen 1912 yılında tanımlanmıştır (Şekil 2).



Şekil 2. Einthoven Modeli [8].

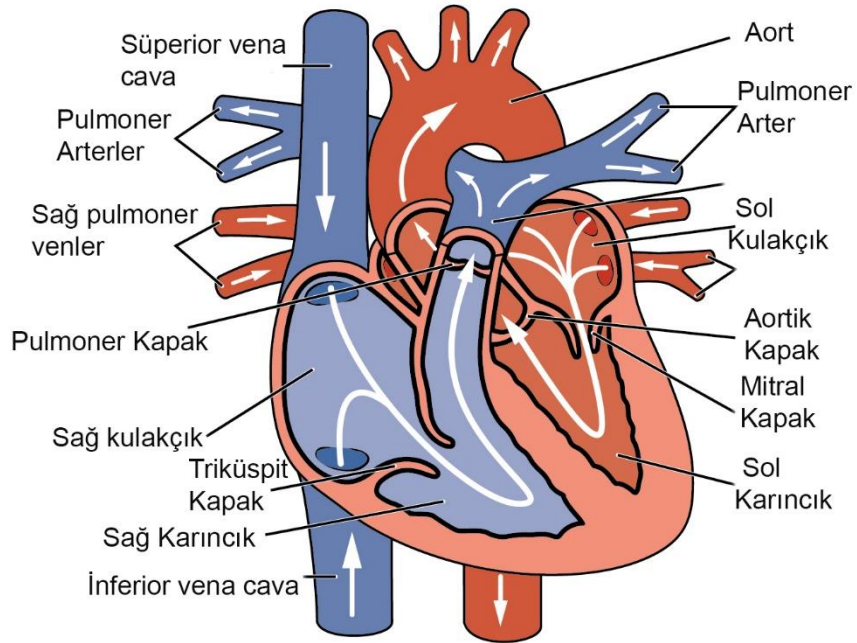
Einthoven modelinde, işaret iki kol ve bir baceden elde edilmiştir. Daha iyi sonuç elde edebilmek için eller ve sol ayak tuzlu su çözeltisi içeren bir kaba daldırılmıştır. Kârlar elektrokardiyografa kablolarla bağlanmıştır. Einthoven 1924 yılında elektrokardiyogram keşfi ile Nobel Ödülü kazanmıştır.

1934 yılında Frank Wilson sağ kol, sol kol ve sol baceğe 50 Ohm dirençler bağlayarak farklı bir elektrot dizilimi önermiştir. Bu günümüzde hala kullanılan “Wilson Central Terminal” olarak bilinmektedir [9]. Tek kutuplu ekstremite derivasyonlar olan VR, VL ve VF Wilson tarafından tanımlanmıştır. V1- V6 Göğüs derivasyonları, standart yerleri ve bağlantı şekilleri Amerikan Kalp Derneği (AHA – American Heart Association) ve Büyük Britanya Kalp Derneği (CSGB – Cardiac Society of Great Britain) tarafından 1938 yılında tanımlanmıştır [10]. Günümüzde kullanılan 12-lead ECG, Emanuel Goldberger tarafından Wilson’un unipolar derivasyonlarının voltajları %50 artırılarak aVR, aVL ve aVF tanımlamaları ile 1942 yılında son halini almıştır [9]. 12-lead ECG Lead I, II, III, aVR, aVL, avF ve V1-V6 göğüs derivasyonlarından oluşmaktadır.

1949 yılında fizikçi Norman Jeff Holter yaklaşık 35kg ağırlığında EKG kayıt sistemini geliştirdi. Günümüzde Holter EKG olarak adlandırılan portatif EKG kayıt sistemlerinde boyut büyük oranda düşürülmüştür. Robert Bruve ve çalışma arkadaşları daha sonradan Bruce Protokol olarak adlandırılacak bir egzersiz test protokolü yayınlamışlardır [11].

1.4.2. Kalbin Anatomik Yapısı

Kalp Şekil 3’de görüldüğü gibi bir pompa görevi göreyerek kanın vücut içerisinde dolaşımını sağlayan dört odacığa sahiptir. Sağ taraf ve sol tarafında birer tane kulakçık ve karıncık vardır. Üst odacıklara (atriyum) sağ kulakçık ve sol kulakçık adı verilir ve bunlar kan toplayıp geçici olarak tutma ve alt odacıklara gönderme görevini yapmaktadır. Alt odacıklar (karıncıklar – ventrikül) kalbin ana pompalama odaları olarak çalışır ve akciğer ve vücudun geri kalanına kanı pompalamaktan sorumludur [1]. Pompalama işlemi diyastol ve sistol olmak üzere iki aşamalı bir işlemdir. Sistol kasılma ve pompalama aşamasıdır. Kulakçık ve karıncıklardaki kasılmalar elektriksel bir tetikleme tarafından düzenlenir. Bu elektriksel tetiklemeler kalbin yapısı üzerinden yayılır ve kalp kası hücrelerinin depolarizasyon ve repolarizasyonunu sağlarlar.

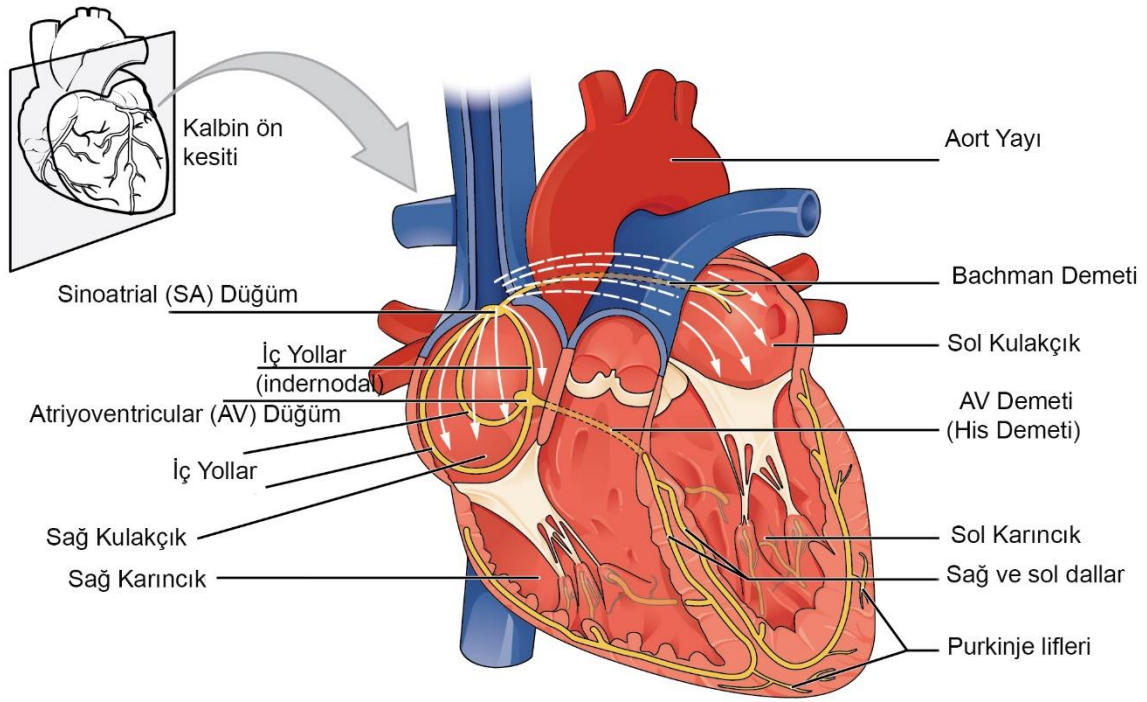


Şekil 3. Kalbin anatomik yapısı [1].

1.4.3. Kalbin İletim Sistemi

Normal bir kalp ritminde, kalbi düzenli bir şekilde tetikleyen tetikleme sinyali üreten ve bu sinyali kalp hücrelerine taşıyan özel bir iletim sistemi vardır. Tetikleme işlemi, kalbin

tetikleme özelliği olan hücreleri (pacemaker) olarak bilinen sino-atriyal (SA) düğümünden başlar, kalp iletim sistemi kullanılarak diğer hücelere dağıtılır. Bu özelliği ile SA düğümü kalbin atım hızını ve ritmini kontrol eder. SA düğümü sağ kulakçıkta bulunur ve buradan başlayan uyarı ilk olarak kulakçıkları uyarır. Tetikleme ara yollar (internodal) üzerinden hem kulakçıklara hem de atriyo-ventriküler (AV) düğümüne iletilir. Tetikleme sırasıyla AV düğümünden his demetine, sağ ve sol dallara ulaşır. Dalların uç kısmında bulunan purkinje lifleri ile karıncıklara yayılır [5]. Şekil 4'te normal sinüs ritminin tetikleme sırası gösterilmektedir.



Şekil 4. Kalbin iletim sistemi [1].

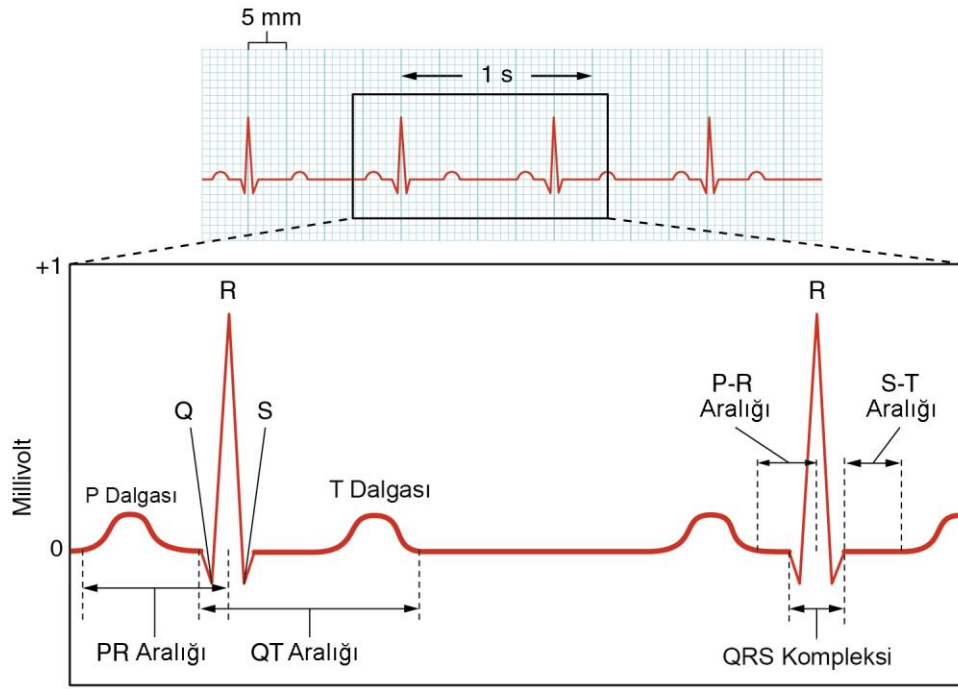
Kalbin normal bir vuru döngüsündeki olayların sırası aşağıdaki gibi gerçekleşir (Şekil 4-6).

- SA düğümü ilk tetiklemeyi gerçekleştirir.
- Elektriksel tetikleme nispeten daha düşük oranlarda atriyal kas yoluyla yayılır ve kulakçıklarda depolarizasyona (kasılmaya) sebep olur. Bu Şekil 5'te görülen EKG dalgasındaki P dalgasını oluşturur. P Dalgası kulakçıkların boyutunun küçük olması

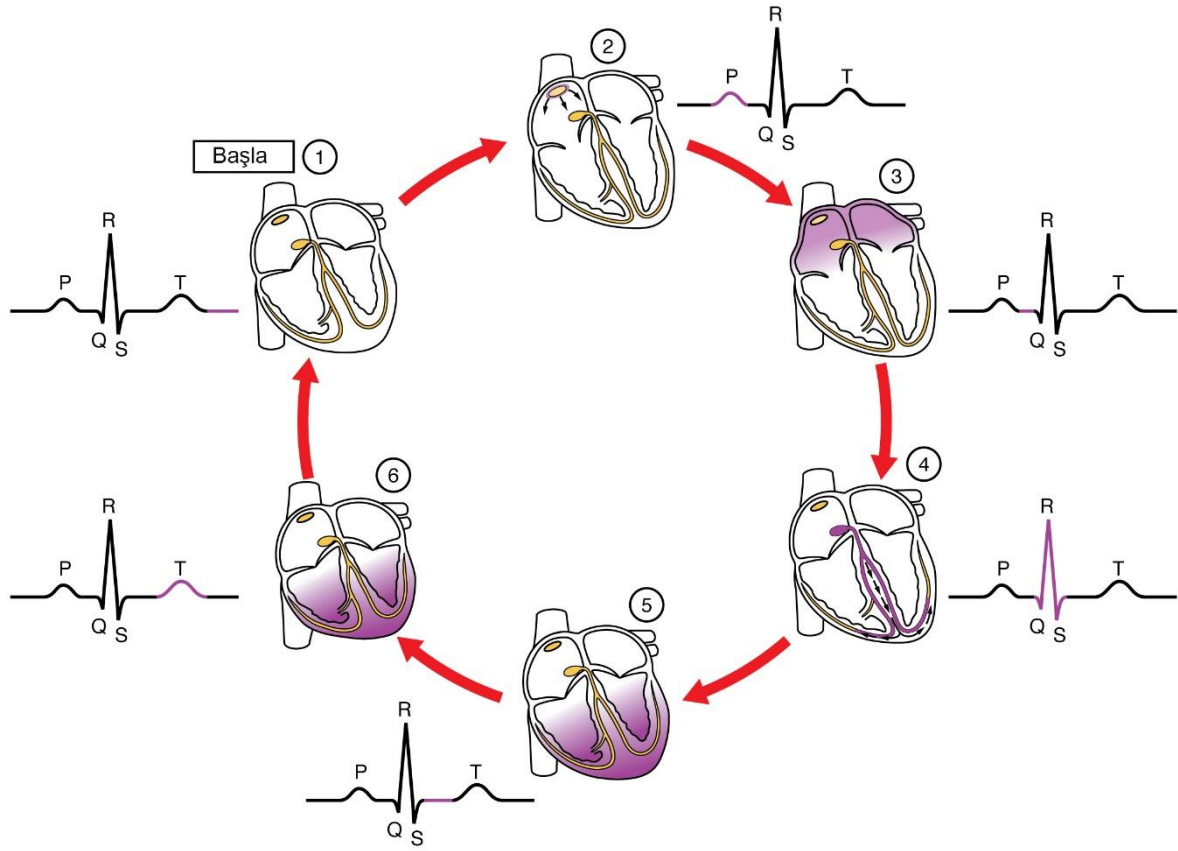
ve yavaş kasılmadan dolayı, yavaş ve düşük yüksekliği olan bir dalgadır. 0,1 – 0,2 mV yüksekliğinde ve 60 – 80 ms arasında bir süre olarak ölçülmektedir.

- Tetikleme dalgası AV düğümünde bir yayılma gecikmesi ile karşılaşır. PQ aralığı olarak bilinen bu gecikme ve P dalgasından sonra normal izo-elektrik seviyesinde yaklaşık olarak 60 – 80ms'dır. Duraklama kanın kulakçıklardan karıncıklara aktarılmasının tamamlanmasına yardımcı olur.
- His demeti, dalları ve purkinje lifleri uyarıyı karıncıklara yüksek hızla iletir.
- Uyarı dalgası kalbin apeksinden yukarıya hızla yayılarak karıncıkların hızla depolarizasyonuna (kasılma) sebep olur. Bu Şekil 5'te görüldüğü gibi EKG'deki 1mV yükseklik ve 80ms süreli QRS dalgasını oluşturur.
- Ventriküler (karıncığa ait) kas hücreleri 300 – 350ms'lik nispeten uzun bir aksiyon potansiyel süresine sahiptir. Aksiyon potansiyelinin QRS dalgasından sonra gelen düz kısmı ST aralığı olarak bilinir ve normal izo-elektrik seviyesinde 100-120ms sürer.
- Karıncıkların repolarizasyonu (gevşemesi) 120-160ms süresi ve 0,1 – 0,3mV genliği ile yavaş bir T dalgasına sebebiyet verir.

Kalbin ritmik normal aktivitesindeki herhangi bir bozukluk aritmi olarak adlandırılır. Kardiyak aritmeye SA düğümünden çıkan düzensiz bir tetikleme formu veya kalbin diğer bölümlerinden çıkacak anormal ve ilave tetikleme aktivitesi sebep olabilir [6].



Şekil 5. EKG dalga şekli ve üzerindeki önemli dalgalar [1].



Şekil 6. Kardiyak döngüsü ve EKG ilişkisi [1].

1.4.4.EKG İşaretinin Karakteristik Özellikleri

Frekans domeninde EKG işaretlerinin enerjisinin çoğu 0,5-40 Hz arasında yer alır [12]. EKG işaretlerindeki morfolojik değişimler genellikle 100 Hz altındaki frekanslarda bulunmaktadır. Bu nedenle EKG kayıtlarının 100Hz ile sınırlandırılması veya kaydedilmesi analizler için yeterli olacaktır [13]. Bu aralıktaki EKG işaretlerini incelemek için 250-360 Hz yaygın örnekleme oranı olarak kullanılmıştır. Çalışmada kullanılan MIT-BIH Aritmi Veritabanında örnekleme oranı 360 Hz'dir [14].

EKG işaretlerindeki gürültü biyolojik ve çevresel gürültü olarak iki ayrı kategoriye ayrılabilir [15]. Biyolojik gürültüler taban gezintisi, hareketten oluşan kas gürültüleri veya elektromiyogram (EMG), nefes almadan dolayı gerçekleşen EKG voltajında görülen değişim, elektrotların hareketi sonucu elektrot-deri direncindeki değişimden kaynaklanan gürültüler olarak bilinir. Çevresel gürültülerin en yaygınları ise kayıt cihazının elektrik hattından kaynaklanan gürültüdür. Diğer çevresel gürültü kaynakları ise EKG kayıt cihazı gürültüsü, diğer elektronik cihazlardan oluşan gürültüler ve elektrot kontak gürültüsüdür [13,15].

Taban Hattı Sapması

Taban hattı sapması EKG işaretini de içeren biyolojik işaretlerin en yaygın gürültüsüdür. Hastanın nefes almasına, hareket etmesine ve harekete dayalı elektrotların değişmesinden kaynaklanabilir. Dinlenme anında kaydedilen EKG'de taban hattı sapmasının ana sebebi hastanın nefes almasıdır. Ancak hastanın hareket etmesi veya nefes almadan dolayı elektrotların değişimi özellikle stres testi veya egzersiz EKG kaydında daha yaygındır [16–22]. Taban hattı sapması için literatürde önerilen birkaç frekans aralığı vardır [16]. Ancak bunlardan en yaygını 0.5Hz'in altıdır. Taban hattı sapmasını genliği, işaretin tepeden tepeye genliğinin %15 olarak ölçülmüştür [23–26].

Elektromiyogram Gürültüsü

EMG gürültüsü milivolt seviyesindeki kas kasılmaları sonucu ortaya çıkar. EMG gürültüsü sıfır ortalamalı bant-sınırlı Gauss gürültüsü gibi işlenebilir. Beklenen genliği

tepeden tepeye genliđin %10'dur ve yaklaşık 50ms uzunluđundadır. Frekans bileşenleri olarak 0 ile 10.000Hz arasında olabilir. EMG işaretlerinin çođu 20-200 Hz aralıđındadır [23].

Güç Hattı Sapması

Güç hattı sapması gürültüsü 50Hz ve harmoniklerinden meydana gelmektedir. Bu gürültüyü oluşturan temel frekansa cihazların güç kaynađında kullanılan alternatif akım sebep olmaktadır. Genliđi tepeden tepeye ECG genliđinin %50'sine kadar olabilir [13,23,27].

Elektrot Kontak Gürültüsü

Elektrot kontak gürültüsü elektrot ile derinin temasının kaybolmasından oluşur [23]. Hastanın hareketine bađlı olarak elektrotların deri ile teması kısa süreli kopabilir. Elektrotun bu hareketi EKG'de ciddi gürültüye sebebiyet verebilir. Bađlantının koptuđu zamanlarda deriden gelen bir akım olmadığı için işaretde güç hattı sapması belirgin olur.

Hareketten Kaynaklanan Gürültüler

Hareketten kaynaklanan gürültüler elektrot ile deri arasındaki direncin deđişiminden kaynaklanan taban hattı deđişimleridir. Bu genellikle verilerin kaydedilmesi esnasında hastanın hareket etmesi ile tetiklenir. Harekete bađlı gürültülerin genliđi tepeden tepeye EKG genliđinin 5 katı kadar olabilir. Genellikle süresi 100-500ms arasındadır [23].

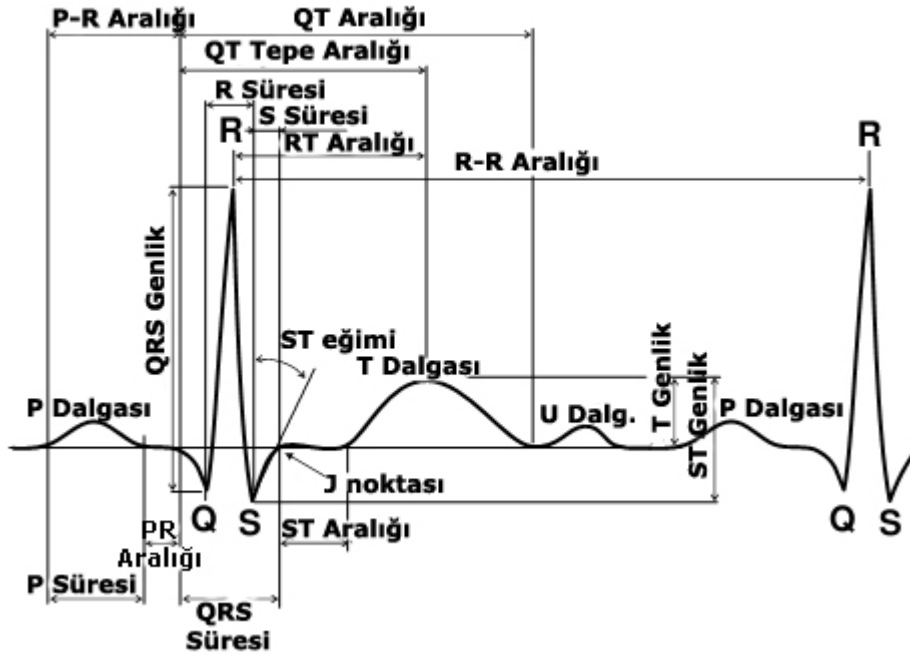
1.4.5. EKG'ye Genel Bir Bakış

EKG'nin iki ana tipi vardır: Standart (dinlenme durumunda kaydedilen) EKG ve eforlu EKG. Her iki EKG vücut üzerinden kayıt edilir. Standart dinlenme zamanı kaydedilen EKG daha yaygın olmasına rağmen efor EKG büyük önem taşımaktadır. Efor EKG genellikle şüpheli veya kanıtlanmış kardiyovasküler hastalıđı olan hastaların deđerlendirilmesinde kullanılır. Stres testi sırasında kaydedilen EKG, bir hastalıđın olup olmadığını, kalbin fonksiyonel kapasitesinin düzenini ve koroner arter hastalıđı (KAH) olasılıđını belirlemek

için kullanılır. Stres testi hastanın bir koşu bandı veya egzersiz bisikletinde egzersiz yaparken kalbinin izlenmesini gerektirir. Stres testinin amacı dinlenme durumunda belirgin olmayan anormallikleri egzersiz ile kalbe yapay zorluklar çıkararak tespit etmeye çalışmaktır [28].

EKG işareti kalbin elektriksel aktivitesini gösterir ve işarete ait öznitelikler kalp ritmi, vuru hızı ve daha da önemlisi kalp problemleri hakkında bilgi verir. EKG işaretine ait öznitelikler iki ayrı kategoriye ayrılabilir: morfolojik (şekilsel) öznitelikler (zaman domaininde EKG dalga şekilleri) ve matematiksel öznitelikler. EKG işaretinin şekilsel öznitelikleri Şekil 7’de gösterilmektedir.

Genellikle bir doktor tarafından incelenen şekilsel öznitelikler zaman domenindeki orijinal EKG işaretinden çıkarılır (EKG domenleri sonraki bölümde incelenmiştir). Bu özelliklerdeki değişim hastada miyokard iskemi ve enfarktüsü, ventrikül hipertrofisi ve iletim problemleri gibi kalp-damar hastalıkları ve anormalliklerin olduğunu gösterebilir [6].



Şekil 7. EKG'deki şekilsel öznitelikler

Matematiksel öznitelikler genellikle şekilsel özniteliklere matematiksel yöntemler veya işaret işleme teknikleri uygulayarak elde edilir. Örneğin; işarete ait frekans bileşenleri matematiksel özniteliklerdendir.

EKG işaretinin tanı için önemli olan şekilsel öznitelikleri: EKG işareti üzerinde gözlemlenebilen başlangıç noktası, eğimli noktalar, genlik, uzunluk, şekil, aralık ve P, Q, R, S, T ve U dalgalarının alanlarına bağlıdır [29–32]. Yaygın olarak kullanılan öznitelikler şunlardır: P dalgası, PR aralığı, QRS kompleksinin şekli ve uzunluğu, ST aralığı, QT aralığı ve RR aralığı. Bunlara ek olarak yaygın olarak kullanılan QRS eksenini olarak bilinen başka bir EKG özneliği daha vardır. Bu ön planda ventriküler aktivasyon ortalama yönünü temsil eder. Bu ventriküler aktivasyon sekansında değişikliklerin bir göstergesidir ve miyokardiyal hasarın bir göstergesi olabilir [32]. QRS eksen üç ekstremite derivasyon tarafından belirlenir. Standart EKG için temel şekilsel öznitelikler aşağıda alt başlıklarda tanımlanmıştır. Tablo 1’de tipik bir II derivasyonuna ait normal sağlıklı bir erkek bireye ait 60 bpm kalp atım hızındaki normal sinüs ritim değerleri verilmiştir.

Tablo 1. Dakikada 60 vuru atan sağlıklı erkeğin II derivasyon sinüs ritmine ait EKG özellikleri [13].

Özellik	Normal Değer	Normal Limit
P genişliği	110 ms	±20 ms
PQ/PR aralığı	160 ms	±40 ms
QRS genişliği	100 ms	±20 ms
QT aralığı	400ms	±40 ms
P genliği	0,15mV	±0,05 mV
QRS yüksekliği	1,5 mV	±0,5 mV
ST seviyesi	0 mV	±0.1 mV
T genliği	0,3 mV	±0,2 mV

P Dalgası

P dalgası atriyal depolarizasyonu temsil eder. P dalgaları derivasyon II’de derivasyon I’den daha belirgindir. Normal bir P dalgası genellikle 0,08 – 0,12 saniye uzunluğundadır ve yaklaşık 0,25 mV genliğindedir [5].

PR Aralığı

PR aralığı P dalgasının başlangıcında başlar ve Q dalgasının aşağıya doğru bükülmeye başladığı izoelektrik seviyesinin sonu olan QRS kompleksinin başlangıcında biter. Atriyal

depolarizasyon başlangıcı ile ventriküler depolarizasyonun başlangıcı arasındaki periyodu gösterir. Diğer bir ifadeyle PR aralığı atriyal aktivasyonun başlangıcı ile ventriküler aktivasyonun başlangıcı arasındaki süredir. Bu zaman aralığı atriyal kasılma ile ventriküler kasılmadan hemen önce karıncıkların kan ile dolmasını sağlar. Normal PR aralığı 0,12 ile 0,20 saniye arasındadır [5].

QRS Kompleksi

QRS kompleksi ventriküler depolarizasyonu gösterir ve EKG'nin en dikkat çekici özneliğidir. QRS kompleksi Q dalgasının başlangıcı (P dalgasından sonra gelen ilk aşağı yönlü eğri) J noktası olarak bilinen S dalgasının yukarı yönlü eğrisi ile sonlanır. Herhangi bir extremitte derivasyonunda R veya S dalgasının normal voltajı 0.2mV'dan küçüktür ve QRS kompleksinin uzunluğu 0.04 ile 0.12 saniye arasında değişir [5,33].

ST Aralığı

ST aralığı QRS kompleksinin sonu ile T dalgasının başlangıcı arasındaki alandır. Normal şartlar altında ST aralığı izoelektrik temel çizgisi veya T aralığı seviyesindedir. Ventriküler depolarizasyon sonu ile repolarizasyon başlangıcı arasındaki periyodu gösterir.

QT Aralığı

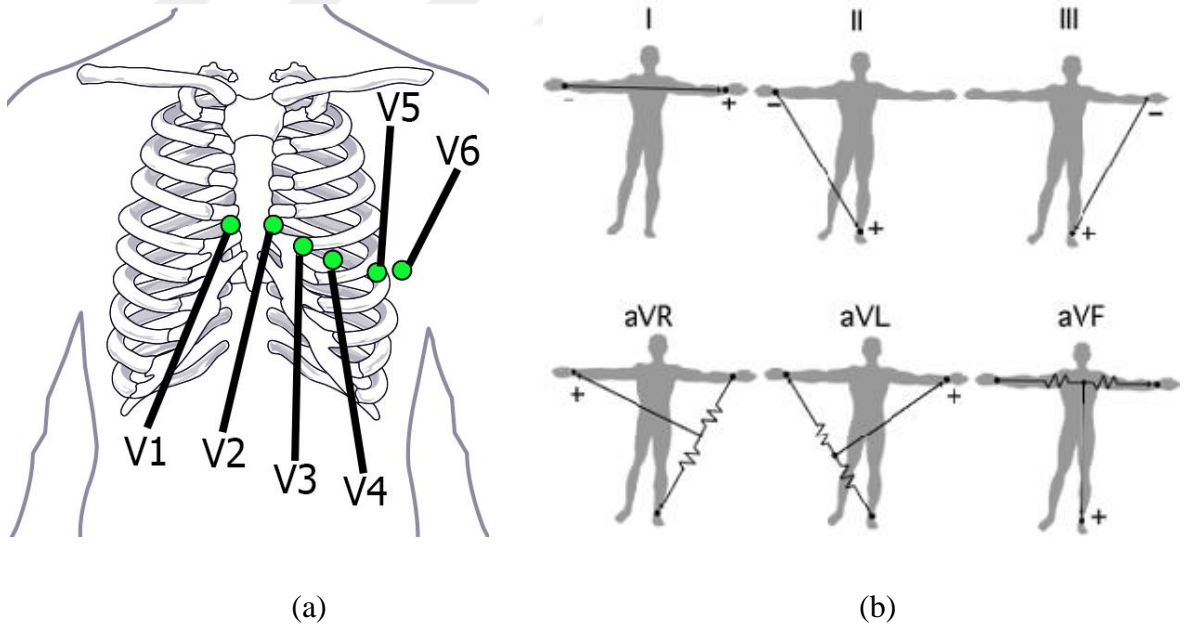
QT aralığı QRS kompleksinin başlangıcından T dalgasının sonuna kadar olan alandır. QT tepe aralığı (Q dalgasının başlangıcından T dalgası maksimum yüksekliği: QTp) bazen QT aralığı yerine kullanılır. Çünkü T dalgasının tepe noktasını algılamak sonunu algılamaktan daha kolaydır [5,33]. QT ile ilgili başka bir ölçü ise düzeltilmiş QT (QTc)'dir. QTc aralığı QT aralığının RR aralığının kare köküne bölünmesi ile elde edilir. QT aralığı karıncıkların depolarizasyon ve repolarizasyon sürelerinin toplamının bir göstergesidir.

RR Aralığı

RR aralığı iki R tepesi arasındaki mesafedir. RR aralığı kalp hızı ile ters orantılıdır. RR arttıkça dakikadaki kalp vuru sayısı azalır. RR aralığı azaldıkça dakikadaki kalp vuru sayısı artar. Normal kalp vuru hızı (HR) 60-90 vuru/dakika aralığındadır. RR aralığı kalp hızı değişkenliğini (HRV) hesaplamak için kullanılabilir. Kalp vuru hızı egzersiz sırasında artar. Dinlenme anında normal bir insana ait kalp vuru hızı sabit değildir, sürekli olarak değişiklik gösterir.

1.4.6. EKG İşaretlerinin Kayıt Edilmesi

Elektrotlar yardımı ile kalbin elektriksel aktivitesinin gözlenebildiği noktalara derivasyon (Lead) denir [5]. Derivasyon teriminin bir diğer anlamı ise elektrotlar tarafından oluşturulan görünüm olarak bilinmektedir.



Şekil 8. Derivasyonlar, (a) Kalp bölgesine ait elektrotların yerleşimi, (b) Standart derivasyonlar.

EKG kaydı için standart 12 derivasyon kullanılmaktadır. Standart 12 derivasyonlu elektrokardiyograf 10 elektrot kullanmaktadır. Derivasyonlardan altısı göğüs

derivasyonlarıdır ve Şekil 8’de görüldüğü gibi kalbin anatomik yerine göre göğsün soluna yerleştirilen 6 tane elektrot ile ölçülür. Geriye kalan 4 elektrot sırasıyla sol kol (LA), sol bacak (LL), sağ kol (RA), sağ bacaktır (RL). Bunlardan sağ bacak elektrotu sistemin topraklaması olarak seçilir. Diğer 6 derivasyon ise 3 tane bipolar ve 3 tane unipolar derivasyonlardır [5,6].

Bipolar derivasyonlar kol ve bacaklara bağlanan elektrotlar arasında tanımlanır. I derivasyonu sol kol ile sağ kol, II derivasyonu sağ kol ile sol bacak, III derivasyonu ise sol bacak ve sol kol arasındaki potansiyel farkı ölçmektedir.

Unipolar derivasyonlar aVL (Sol kol), aVR (Sağ kol), aVF (Sol bacak) olarak bilinirler ve bipolar derivasyonlardaki elektrotlar kullanılarak ölçülür ve genliği artırılarak kaydedilir.

Son 20 yılda fizyolojik işaretleri kayıt eden sistemler geliştirilmiş ve iyileştirilmiştir. Bunlar daha hafif, daha küçük ve işaretleri 48 saate kadar kayıt imkânı sağlayan sistemlerdir. Ayrıca ayakta kayıt sistemi olarak adlandırılan bu sistemler genellikle nadir olarak görülen aritmiler, geçici iskemik olaylar veya sessiz miyokard iskemisi gibi günlük yaşamın stresiyle ilgili kalp fonksiyonlarının geçici anormalliklerini tespit etmek için EKG analizinde kullanılırlar. Bazı hastalık tipleri kısa zamanlı veya 12 derivasyonlu EKG kayıtlarında algılanamayabilir. Bu cihazlar Holter cihazları olarak bilinir ve N. J. Holter tarafından 1961 yılında geliştirilmiş EKG kaydedicileridir [34]. Katı hal bellek depolama o zaman olmadığından ilk Holter kayıt cihazı büyük işaret dosyalarını kaydetmek için kaset kullanmaktaydılar. Şu anda kullanılan Holter cihazlarında işaretler flaş tipi yarı iletken belleklere kaydedilmekte ve üzerinde daha fazla analiz yapılabilmesi için iş istasyonlarına aktarılmaktadır.

Standart klinik EKG’nin bazı önemli özellikleri şunlardır:

- 1mV yükseklik ve 200ms genişlik değerlerinde kâğıt üzerine çizilecek bir vuru için 1cm yüksekliği gösterecek şekilde kalibrasyonu yapılır.
- Kağıt hızı olarak 25mm/s ölçüsü kullanılır. Bu da grafik ölçeği olarak 0.04s/mm veya 40ms/mm olmasını sağlar. Kalibrasyonun vuru genişliği 5mm olur.
- Bir ECG işaretinin tepe değeri normal olarak 1mV civarında olur.

- Klinik EKG genellikle tanı EKG'si için önerilen 500 Hz örnekleme oranı ile 0,05-100 Hz bant genişliğinde filtrelenir. Kalibrasyon vurusunun biçimindeki bozukluk filtre ayarlarının uyumsuzluğu veya işaret alma sisteminin zayıflığını gösterir.
- İzleme için EKG'de 0,5-50 Hz düşürülmüş bant genişliği kullanılabilir.
- Yüksek çözünürlüklü EKG 0,05-500 Hz'den daha yüksek bant genişliklerini gerektirir.

1.4.7. Klinik EKG Analizine Temel Bir Yaklaşım

Klinik elektrokardiyogram analizinde sistematik bir yaklaşım kullanmak önemlidir. Klinik yaklaşımı gösteren aşağıdaki özet kardiyologların EKG'deki anormallikleri nasıl tanımladıklarını anlama açısından basit bir rehber olabilir [13].

- QRS kompleksini tanımlayın. Burada şu gözlemler yapılmalıdır:
 - Ventriküler oran nedir?
 - QRS kompleksleri düzenli aralıklarla mı ilerliyor? Değilse düzensizliğin sebebi nedir?
 - Verilen derivasyondaki QRS kompleksleri aynı mı? Normal boyut ve şekildedir mi?
- P dalgasını tanımlayın. Bazı durumlarda çok dikkatli inceleme gerektirebilir ve birden fazla derivasyon eksenine gerekli olabilir. Şu sorular araştırılmalıdır:
 - P dalgası ve QRS kompleksi arasında birebir ilişki var mı? Değilse tanımlanabilir bir model var mı?
 - PR aralığının normal süresinde mi?
 - Atriyal oran nedir?
 - Verilen derivasyonda P dalgaları şekilsel olarak tanımlanabiliyor mu? Normal boyut ve şekildedir mi?

Bu iki madde ile çoğu durumda ritim mekanizmasını tanımlamak mümkün olmaktadır.

- Her derivasyondaki QRS kompleksini inceleyin. QRS eksenine normal mi? Genel olarak QRS genişlik ve genlikleri normal mi? Genellikle QRS kompleksleri kalbin belli bölümlerine özgü gruplar halinde görüntülenir. Dalga desenleri kalple ilgili

ara iletim blokları ve benzeri problemlerin tespiti için kontrol edilmeli. Q dalgasındaki önemli şekil bozuklukları ve R dalgasındaki değişiklikler önemlidir.

- ST-T aralığını inceleyin. Anormallik var mı? Anormallik iskemi, enfarktüs ve hipotermi andırıyor mu?
- T dalgalarını inceleyin. Şekilleri normal mi? Her derivasyonda QRS kompleksi ile aynı yönde yönlendirilmiş mi? Değilse iskemi veya ventriküler ileti anormalliklerini andırıyor mu?
- QT aralığını inceleyin. RR aralığının yarısında bitti mi [13]?

1.4.8. Kardiyak Bozukluklara Genel Bir Bakış

EKG kullanılarak gözlenen patolojiler üç kategoriye ayrılır [35]:

- a) Kalp ritim bozuklukları veya aritmiler.
- b) Miyokardaki veya kardiyak iskemideki kan akımında işlev bozukluğu.
- c) Sol ventrikül hipertrofisi gibi kalbin mekanik yapısındaki kronik hastalıklar.

Bu çalışmada yukarıda tanımlanan a tipi patolojilere odaklanmıştır. Özellikle kalbin ventriküler elektrik aktivitesini gösteren QRS kompleksi üzerinde deneyler gerçekleştirilmiştir. QRS kompleksindeki önemli değişimler anormal ventriküler kalp vuruları hakkında değerli klinik bilgiler içermektedir. QRS kompleksi kalp vurularının çoğunda mevcuttur ve tüm dalgalar içerisinde işaret gürültü oranı en yüksek olanıdır.

EKG analizindeki temel sorun işaret morfolojisindeki geniş çeşitliliktir. Sadece hastadan hastaya değişen özellikler değil, hastanın hareketi, elektrik iletimindeki değişimler, vücut karakteristikleri gibi özellikler EKG işareti üzerinde şekilsel değişikliklere sebep olabilmektedir. Bundan dolayı, ilgilenilen tüm durumları dikkate alan bir eğitim seti hazırlayabilmek mümkün değildir. Buna ek olarak EKG işareti biyolojik kaynaklar ve harici kaynaklar gibi gürültü kaynakları tarafından bozulmuş olabilmektedir. Sınıflandırıcının kullanacağı giriş işareti performansını doğrudan etkileyeceği için bu tür çalışmalarda işaret ön işleme için uygun tekniklerin seçilmesi çok önemlidir.

Aritmiler iki gruba ayrılabilir. Birinci grup hayati tehdit oluşturan ve hemen tedavi gerektiren ventriküler fibrilasyon ve taşikardiden oluşmaktadır. İkinci grup ise acil hayati tehdit oluşturmayan ama tedavisi ile ileride oluşabilecek problemleri önleyen aritmileri içermektedir [35].

AAMI (ANSI/AAMI EC57:1998/(R)2003) standartlarında bulunan gruplar ve MIT-BIH aritmi veri tabanındaki gruplar Tablo 2’de gösterilmektedir [36,37]. MIT-BIH Aritmi Veritabanı [37] kardiyak aritmilerin analizi ile ilgili algoritmaların tasarımını değerlendirmek için kullanılan en yaygın veritabanıdır. Veritabanı AAMI tarafından önerilen her bir aritmi grubu içerisinde birçok vuru türünden oluşmaktadır. Örneğin Normal grupta LBBB aritmi tipi, RBBB aritmi tipi, AE ve NE aritmi tipleri bulunmaktadır.

Tablo 2. AAMI standartlarına göre vuru tipleri

AAMI Kalp Vurusu	Tanım	MIT-BIH Vuru Tipleri
N	Sv, V, F ve Q dışında kalan tüm vurular	Normal (N), Sol dal blok vurusu (LBBB), sağ dal blok vurusu (RBBB), Kulakçık Kaçak vurusu (AE), Nodal (junctional) kaçış vurusu (NE)
Sv	Supraventriküler ektopik vuru	Kulakçık Erken Vurusu (AP), Anormal Kulakçık Erken Vurusu (aAP), Dügümsel Erken Vuru (NP), Supraventriküler Erken Vuru (SP)
V	Ventriküler ektopik vuru	Erken Karıncık Kasılması (PVC), Karıncık Kaçak Vurusu (VE)
F	Füzyon vuru	Normal ve Karıncık Vurusu birleşimi (fVN), Normal ve Yapay Vuru birleşimi (fPN)
Q	Tanımlanmayan vuru	Yapay Vuru (P), Sınıflandırılmamış Vuru (Q)

Kalple ilgili önemli hastalıklardan bazıları kalbin anatomik yapısındaki anormallikler sonucunda değil, anormal ritimleri sonucunda ortaya çıkmaktadır. Kalbin normal fizyolojik davranışlarından sapmasına aritmi denir. Aritmiler genellikle anormal pompalama

fonksiyonu ile ilişkilidir ve hayat kalitesinin düşmesine hatta ölüme sebebiyet vermektedir [38].

Aritminin birçok türü vardır. Aritmiler kalbin hangi bölgesinde ortaya çıktıkları (karıncık veya kulakçık) ve meydana geldiklerinde kalbin ritminde ne gibi değişiklik olduğuyla tanımlanırlar. Ektopik vuru veya desen tipli aritmiler olarak da sınıflandırılabilirler.

Ektopik kalp vuruları kalp atımlarındaki düzensizliklerdir ve fazladan yapılmış veya eksik kalp vurularını içeren kalp ritimleridir. Ektopik vuru olarak bilinen ekstra kalp vuruları çok yaygın hastalıklardır. Kalbin karıncık veya kulakçıklarında meydana gelebilir. Ektopik vurular tek başına tehlikeli değildirler ve kalbe zarar vermezler. Ektopik vuru türleri ve özellikleri aşağıda özetlenmiştir.

Supraventriküler ektopik vuru: Karıncık seviyesinin üstünde bir yerde ektopik bir darbenin (dürtünün) neden olduğu bir kalp vurusudur.

Erken kulakçık kasılması: Kalp hızı normal kalır, ancak erken P dalgasından dolayı kalp ritmi düzensizleşir. Bu aritmi tipi çarpıntı, atriyal çarpıntı veya atriyal fibrilasyona neden olabilir.

Atriyal kaçak vuru: Uzun sinüs duraklamalarının hemen ardından ortaya çıkan atriyal ektopik vurulardır. Tek veya çoklu olabilirler. Kaçak vurular ektopik atriyal ritim olarak adlandırılan sürekli bir ritim üretebilir. Kalp vuru hızı genellikle daha yavaştır, P dalga şekli farklıdır ve PR aralığı normal sinüs ritminden biraz daha kısadır.

Karıncık erken vurusu: Normal bir kalp vurusu oluşmadan hemen önce oluşan, karıncıklardan kaynaklanan anormal bir elektriksel aktivite sonucunda ortaya çıkan ekstra bir kalp vurusudur.

Erken karıncık kasılması: Kalp hızı değişkendir. P dalgası genellikle erken karıncık kasılmasının QRS, PST veya T dalgası tarafından örtülmüştür. QRS kompleksinin genişliği 0.12sn'den daha fazladır ve şekilsel olarak alışılmadık bir biçimde ST aralığı ve T dalgası ile zıt kutupludur. QRS kompleksi çok odaklı ve farklı morfolojileri gösterebilir.

Karıncık kaçak vurusu: Bir ritimdeki uzun bir aradan sonra meydana gelen ektopik bir vurudur. SA düğümünün bir vuruyu başlatmasındaki hatasından veya bu vurunun AV düğümüne iletilmesindeki hatadan dolayı oluşmaktadır.

Erken bileşke vuru: AV düğüm birleşimine yakın bir yerde ortaya çıkar. Genellikle herhangi bir tedavi gerektirmez.

Sol dal bloğu: Sol karıncığın aktivasyonu gecikir, bu da sol karıncığın sağ karıncıktan daha sonra kasılmasına sebep olur. QRS kompleksinin süresinin uzamasına neden olur.

Sağ dal bloğu: Sağ dal bloğu sırasında sağ karıncıklar sağ dal üzerinden gelen dalgalar tarafından doğrudan aktifleştirilmez. Ancak sol karıncık hala normal sol dal tarafından aktifleştirilir ve bu dalgalar sol karıncığın miyokardiyumu üzerinden geçerek sağ karıncığa ulaşır ve sağ karıncığı aktifleştirir. Bu bekleme QRS kompleksinin genişlemesine sebep verir.

Bileşke kaçak vuru: AV birleşimi civarında oluşan bir tetikleme tarafından üretilen gecikmiş bir kalp vurusudur. SA düğümünün depolarizasyon oranı AV düğümünün oranının altına düştüğünde gerçekleşir. SA veya AV bloku olduğundan dolayı SA düğümünden çıkan elektriksel darbe AV düğümüne ulaşamadığı durumda da bu ritim bozukluğu meydana gelebilir.

Diğer tür aritmiler desen tipi aritmilerdir. Bu türden aritmiler birbirini takip eden vuruların karakteristikleri ile tanınırlar ve supraventriküler veya ventriküler olarak gruplandırılırlar. Supraventriküler aritmiler kalbin üst iki odacığında (atriumlarda) meydana gelir. Supraventriküler aritmi tipleri atriyal fibrilasyon (AF), atriyal flutter, paroksizmal supraventriküler taşikardiden meydana gelmektedir. Ventriküler aritmiler ise kalbin iki alt odacığında (karıncık) meydana gelir. Ventriküler aritmi tipleri ventriküler fibrilasyon (VF) ventriküler flutter ve ventriküler taşikardiden meydana gelir. Bu aritmilerin en tehlikelisi ventriküler aritmilerdir, çünkü ölüme sebebiyet verebilir [38].

Atriyal fibrilasyon: Bu elektriksel bir ritim bozukluğudur. Atriya da meydana gelen anormal elektriksel tetikleme sonucunda kasların hatalı olarak kasılması ve kanın verimsiz olarak pompalanmasına neden olur. Bu nedenle üst odacıklarda (atrium) bulunan kan tamamen alt odacıklara boşaltılamaz. Atriallarda kanın beklemesi kırmızı kan hücrelerinin

birbirine yapışmasına ve kan pıhtılaşmasına neden olabilir. Atriyal fibrilasyonun en tehlikeli sonucu pıhtının kalpten çıkarak yer değiştirmesi ve pıhtının vücudun en önemli organlarından birinde damar tıkanıklığına yol açmasıdır [38,39].

Ventriküler fibrilasyon: Ventriküler fibrilasyon farklı ventriküler odaklardan birçok düzensiz ve koordinesiz tekrar eden hızlı uyarıların çıkması ile oluşur. Ventriküler fibrilasyon EKG’de rasgele ve birbiri ile ilişkisiz dalgalara sebep olur. Birçok ventriküler noktadan aynı anda çıkan uyarılar EKG’de düzensiz ince salınımlar oluşturur. Ventriküler fibrilasyon ölümcüldür, çünkü ritim tümüyle düzensizdir ve karıncıklar kanı etkin bir biçimde pompalayamaz.

Ventriküler flutter: Kalbin dakikadaki vuru sayısı 250’yi geçtiğinde tehlike arz eder. Bu tür aritmide kalp odacıkları o kadar hızlı kasılır ki kanın akması ve odacıkları doldurması için zaman yoktur. Bu durumda kalp dolaşım sistemine çok az miktarda bir kan transfer eder. Bu durumla karşı karşıya olan bir kimse kısa süre içerisinde bilincini kaybeder.

Ventriküler Taşikardi: Ventriküler taşikardi ventrikül içerisinde başlayan 3 veya daha fazla ardışık erken ventrikül atımdan tanımlanan hızlı kalp atımlarıdır.

1.5. EKG İşaretlerinin İşlenmesi

1.5.1. Domenler

ECG işaretlerinde yaygın olarak kullanılan domenler ilk olarak filtreler, dalga algılama yöntemleri ve özellik çıkarımı teknikleri gibi EKG işaret işleme teknikleri tarafından tanımlanmıştır. Bunlar zaman, frekans ve dalgacık domenidir.

Zaman Domeninde EKG İşaretleri

Zaman domeninde EKG işaretleri genlik ve süreleri gibi işaretin özelliklerini temsil eder [40]. Örneğin, P dalgası ve QRS kompleksi analizler için genlik (mV) ve süre (ms) olarak ifade edilir. Zaman domeni öznelik çıkarımı ve analizi, önemli EKG işaret analizi yöntemlerindedir. Daha önceki bölümlerde bahsedildiği gibi kardiyak bozukluklar EKG

işaretine bileşenlerdeki değişim olarak yansır. İncelenecek tipik EKG bileşenleri P dalgası, R dalgası, QRS kompleksi, ST segmenti ve T dalgasıdır. Buna ek olarak PR aralığı, RT aralığı, QT aralığı, TP aralığı ve RR aralığı sıklıkla incelenmiştir. RR aralığından türetilen kalp hızı değişkenliği (heart rate variability - HRV), tipik bir EKG zaman domeni analizi yöntemidir.

Frekans Domeninde EKG İşaretleri

EKG işaretleri zaman domeninde kaydedilir ve işaretin karakteristikleri genlik ve süreleri olarak temsil edilmektedir. Bununla birlikte, işaret içerisinde mevcut frekanslar aynı zamanda işaretleri anlamak ve analiz etmek için kullanılabilir. EKG işaretindeki enerjinin çoğu 0,5 – 40 Hz frekansları arasındadır [12,41]. EKG işaretleri baskın morfolojik özellikler genellikle 100 Hz altında frekansları içerir. Bu aralıktaki işaretlerin incelenmesi için yaklaşık bir örnekleme oranı yeterli olacaktır (250 veya 360 Hz). Bu alanda yapılmış çalışmalarda kullanılan en yaygın EKG veritabanı olan MIT-BIH Aritmi Veritabanı bu aralıkta bir örnekleme frekansına sahiptir [14].

EKG işaretinin frekans domen bilgisi genellikle gürültü giderme işlemlerinde kullanılan filtreleme gibi ön işlemler için kullanılmaktadır [15,17,18,42–52]. Frekans domeninde özellik çıkarımı literatürde yaygın olarak kullanılmaktadır.

ECG işaretlerinde genellikle 100 Hz altındaki frekans bileşenleri analiz edilmesine rağmen daha yüksek frekans bileşenleri de çalışılmıştır [53]. Anlamlı bilgi içeren frekans üst limiti tam olarak bilinmemekle birlikte 150-250 Hz aralığı en yaygın çalışılmış frekans aralığıdır [54–57]. Yüksek frekanstaki problemlerden biri genliğin çok düşük olmasıdır. Bu da gürültüden sıklıkla etkilenmektedir.

Dalgacık Domeninde EKG İşaretleri

Dalgacık domeni daha önce bahsedilen domenlere göre nispeten daha yeni bir yaklaşımdır. EKG işaretleri bu domende işlenmektedir [58–68]. Dalgacık domenindeki işaretin bilgisi zaman domenindeki işarete dalgacık dönüşümü uygulayarak elde edilebilir. Dalgacık dönüşümü ile zaman domenindeki işaret, alçak geçiren ve yüksek geçiren filtrelere

benzeyen *yaklaşım* olarak bilinen alçak frekans bileşenlerine ve *detay* olarak bilinen yüksek frekans bileşenlerine ayrıştırılır [69]. Bu bağlamda dalgacık domenindeki analiz frekans domenindeki analize oldukça benzer. Aslında dalgacık analizi Fourier dönüşümünün aksine zaman ve frekans domen bilgisinin her ikisini de sağlar [70]. Bu dalgacık domenindeki bilginin en temel özelliklerinden bir tanesidir.

Dalgacıkların bir diğer önemli özelliği ise çok çözünürlülük olarak bilinir. Daha önce bahsedilen detay ve ayrıntı kavramları biraz daha ayrıştırılabilir. Dalgacık dönüşümü çok çözünürlülük altyapısı sayesinde EKG işaretlerinin işlenmesi için en güçlü araçlardan bir tanesi haline gelmiştir [71]. Dalgacık analizi EKG işaretlerinde filtreleme [48,49,52,72–74] ve sıkıştırmayı [46,75–83] içeren ön işleme ve özellik çıkarımında [59,60,84,85] yaygın olarak kullanılmıştır.

1.5.2. R Tepesinin Algılanmasında Kullanılan Algoritmalar

QRS kompleksi karıncıkların elektriksek aktivitesini göstermektedir. Bu EKG işaretindeki en öne çıkan özelliktir. R tepesi ise QRS kompleksinin en ayırt edici bileşenidir. Bu yüzden ilk olarak R tepesinin tespit edilmesi gerekir [86]. R tepesi ayırt edici dalga şeklinden dolayı EKG işaretlerinin sınıflandırılmasında, kardiyak döngüsünün belirlenmesi için kullanılır [87]. QRS kompleksi tanımlandığında diğer bileşenlerin algılanması ve EKG işaretinin incelenmesi gerçekleştirilebilir. Örneğin ST segmenti QRS kompleksinin sonunda bulunmasından dolayı QRS'nin tespitinden hemen sonra kolaylıkla tanımlanabilir [88]. P dalgası tespit edilmeden önce R tepesi ve QRS kompleksinin ilk olarak bulunması gerektiği yapılan çalışmalarda belirtilmiştir [89].

Son yıllardaki çalışmalarda R tepesinin algılanması için birçok algoritma geliştirilmiştir. Bu algoritmalar genellikle iki aşamadan oluşmaktadır: Öznitelik çıkarımı ve sınıflandırma [61,84,90–92]. R tepesi öznitelik çıkarımı genellikle filtreleme veya gerekli olmayan bilgilerin bastırılması için işaretin ayrıştırılması işlemlerini içerir. R tepesi algılama algoritmalarının çoğunda temel tanıma öncesi gürültü temizleme için yüksek-geçiren filtre veya alçak geçiren filtre kullanılmaktadır [86]. R tepesinin çıkarılan öznitelikleri karşılaştırılırken ön tanımlı veya adaptif eşikleme en yaygın kullanılan yöntemlerdendir.

Literatürde yapılan çalışmalarda R tepesinin algılanmasında QRS algılama [93,94] olarak bahsedilmektedir. 1983 yılında Ligtenberg ve Kunt [95] R tepesini 3 kategoride sınıflandırmıştır: eşleşme algoritmaları [96], sentetik algoritmalar [97,98] ve eşik algoritmaları. Friesen 1990 yılında R tepesi algılama tekniklerini 9 algoritma sınıfında kategorize etmiştir [23]. Daha yakın tarihe geldiğimizde ise dalgacık dönüşümü [67,99–106] ve yapay sinir ağları [107–109] R tepesinin algılanmasında yaygın olarak kullanılmıştır. Bu yeni yaklaşımları içeren R tepesinin algılanmasında daha kapsamlı bir çalışma 2002 yılında Kohler tarafından MIT-BIH aritmi veritabanı gibi standart dinlenme anında kaydedilmiş bir veritabanı üzerinde gerçekleştirilmiştir [87]. Kohler'in kullandığı kategoriler: türev ve sayısal filtre tabanlı algoritmalar, dalgacık tabanlı algoritmalar, yapay sinir ağı tabanlı algoritmalar, adaptif filtreler [110], gizli Markov modelleri [111], matematiksel morfolojiler [87], eşleşmeli filtreler [112], genetik algoritmalar, Hilbert-Dönüşümü tabanlı QRS algılama, uzunluk ve enerji dönüşümü, sentetik metotlar [87] QRS algılama algoritmalarıdır. Buna ek olarak Moody ve arkadaşları [113] ve Kaplan [114] QRS kompleksini algılamak için QRS morfolojisini temsil eden temel vektörleri kullanan şablon tabanlı yaklaşımlarını geliştirmişlerdir.

R tepesinin algılanması için yukarıda bahsedilen iki yaygın kavram türev ve sayısal filtreler tabanlı algoritmalar ve dalgacık tabanlı algoritmalar aşağıda ayrıntılarıyla bahsedilmiştir.

Türev ve Dijital Filtre Tabanlı Algoritmalar

Birçok R dalgası algılama algoritması türev ve dijital filtre tabanlı algoritmalar kategorisine aittir [31,115–121]. Bu algoritmaların çoğunda R tepesinin algılanması işlemi için işarete istenmeyen bileşenlerin çıkarılması amacıyla ön işleme yapılmaktadır. Filtrelenmiş işareten çıkarılan öznitelikler belirlenmiş sabit veya adaptif bir eşik değeri ile karşılaştırılarak R tepesi algılanması yapılmaktadır [87].

Dalgacık Tabanlı Algoritmalar

Dalgacık dönüşümü yakın zamanın en çok kullanılan işaret işleme tekniklerinden bir tanesidir. Literatürde R tepesinin algılanmasında birçok uygulaması yapılmıştır [67,91,99–101,103,106,122–124].

Dalgacık dönüşümünün çok çözünürlüklü altyapısı ile bir EKG işareti zamana ilişkin önemli bilgileri koruyarak birçok farklı frekans bantlarına ayrıştırılır [59]. İşaretin R dalgasının güç spektrumu ile ilişkili ayrıştırılmış işaret bantları kullanılarak R tepesi saptanabilir [125]. Bu yöntemle EKG işaretindeki gürültü veya diğer bileşenler gibi tüm gereksiz bilgiler bastırılabilir. R dalgasının tanımlanması için genellikle, ilişkili ayrıştırılmış işaretlerde yerel maksimum belirlenmiş bir eşik değeri ile karşılaştırılır [93].

1.6. EKG İşaretlerinde Öznitelik Çıkarımı

Öznitelik çıkarımı ve öznitelik boyutunun azaltılması temelde büyük miktardaki öznitelik sayılarını daha küçük miktarlardaki öznitelikler ile temsil etmeyi amaçlamaktadır. Boyut azaltma metodlarında iki ana kategori vardır; öznitelik seçimi ve öznitelik çıkarımı. Öznitelik seçimi yöntemlerinde, öznitelik kümesinden sadece yararlı öznitelikleri kullanılıp diğerleri elenir. Öznitelik çıkarma (literatürde öznitelik oluşturma, hesaplama veya dönüştürme olarak bahsedilmektedir) yöntemlerinde ise çeşitli sayısal yöntemler kullanılarak orijinal değişkenlerden yeni öznitelikler hesaplanmaktadır [126]. Öznitelik çıkarma büyük bir veri kümesi tanımlamak için gerekli olan kaynakların miktarını azaltmak için kullanılır. Karmaşık verilerin analizini yaparken önemli sorunlardan biri ilgili değişkenlerin sayısının fazla olmasından kaynaklanmaktadır. Bir analiz esnasında çok fazla değişken kullanımı genellikle büyük bellek gereksinimi ve hesaplama yükü oluşturur. Çok fazla değişken kullanımının bir diğer olumsuz sonucu ise sınıflandırma algoritmasının eğitim setini ezberlemesi ve yeni örnekleri sınıflandırmada zayıf kalmasıdır. Özniteliklerin boyutunu mümkün olduğu kadar düşük tutarak sadece işlem yükü azaltılmaz, aynı zamanda sistemin daha sağlam hale getirilmesi arzu edilir. Örneğin, Chow ve Huang öznitelik boyutu örnek sayısı az bir veri seti kullanan sınıflandırıcının performansının büyük oranda düştüğünü göstermişlerdir [127].

Öznitelik çıkarımı yeterli doğrulukla veriyi tanımlarken problemlerin üstesinden gelebilecek değişken kombinasyonlarını oluşturma yöntemleri için genel bir terimdir. Ham giriş verisi uzayının boyutunun azaltılması örüntü tanıma görevleri için önemli bir adımdır. Boyut azaltmada amaç, veriyi düşük boyutlu uzaya dönüştürerek daha gelişmiş bir sınıflandırma temsili oluşturmaktır. Öznitelikler orijinal işarete uygulanan işaret işleme yöntemine göre farklı domenlerde ifade edilebilir.

EKG işaretinin tipik özellikleri daha önceki bölümlerde bahsedildiği gibi üç temel domende bulunmaktadır. Bunlar zaman, frekans ve dalgacık domenleridir. Aşağıda ayrıntılı olarak ele alınmıştır. 1.6.2 başlığı altında bahsedildiği gibi ilk olarak R tepesinin algılanması gerçekleştirilir. Diğer önemli noktaların algılanması ve tanınması bu işlemden sonra gerçekleştirilir. R tepesinin algılanmasından sonra kalp atım hızı hesaplanabilir. R-R aralığı iki R dalgası arasındaki mesafeyi süre olarak göstermektedir. Kalp hızı dakikadaki vuru olarak ifade edilir. Burada bir vuru bir R-R aralığının mili saniye cinsinden değerini gösterir. Eşitlik (1)'den görüldüğü gibi kalp vuru hızı (KH) ve R-R aralığı ters orantılıdır. EKG işaretlerinin zaman domeninde yaygın olarak çıkarılan öznitelikleri aşağıda kısaca bahsedilmiştir.

1.6.1. Zaman Domeninde Öznitelik Çıkarımı ve Analizi

EKG'nin zaman domainindeki öznitelikleri işaret işleme teknikleri [67,87,109,119,128–134] kullanılarak otomatik olarak veya basit ölçüm yöntemleri ile elde edilebilir.

$$KH = \frac{60}{R - R \text{ aralığı}} \quad (1)$$

Kalp hızı değişkenliği (KHD) otonom sinir sistemi tarafından uygulanan kalp hızının kontrol ölçüsünü gösterir [135].

EKG zaman domenindeki diğer özellikleri ise P dalgası, PR aralığı, QRS kompleksi, ST aralığı, QT aralığı, RT aralığı, PP aralığıdır. P dalgası R'den önce gelen tepe, T dalgası da R'den sonraki tepe olarak tanımlanabilir. Ancak genliklerinin düşük olmasından dolayı P ve T dalgasının algılanması zor olabilir [136].

Zaman domenindeki EKG öznitelikleri iki grupta incelenebilir [137]. Her bir dalganın şeklini tanımlayan öznitelik, örneğin P dalgasının genlik ve genişliği ve önemli noktalar arasındaki aralıklar, örneğin PR veya QT aralıkları. Bu dalgaların şekillerindeki değişimin algılanması gürültü tarafından bozulmuş EKG işaretlerinde zordur. Nispeten temiz ve durağan EKG işaretleri EKG şekil analizi için gerekli olabilir. Yerel olarak tepeden tepeye aralıkların hesaplanması için arama aralığındaki yerel maksimumlar kullanılabilir [138].

EKG işaretinin zaman serisinden tanımlanmış özel bir pencere genişliği ve karşılık gelen işaretin genlik değerleri öznitelik olarak kullanılabilir. Bu en basit öznitelik çıkarma yöntemlerinden bir tanesidir. Herhangi bir ek işlem yüküne gerek duyulmayacaktır. Pencere genişliği burada daha iyi performansı verecek sonucu elde etmek için araştırılacak parametrelerden bir tanesidir [38].

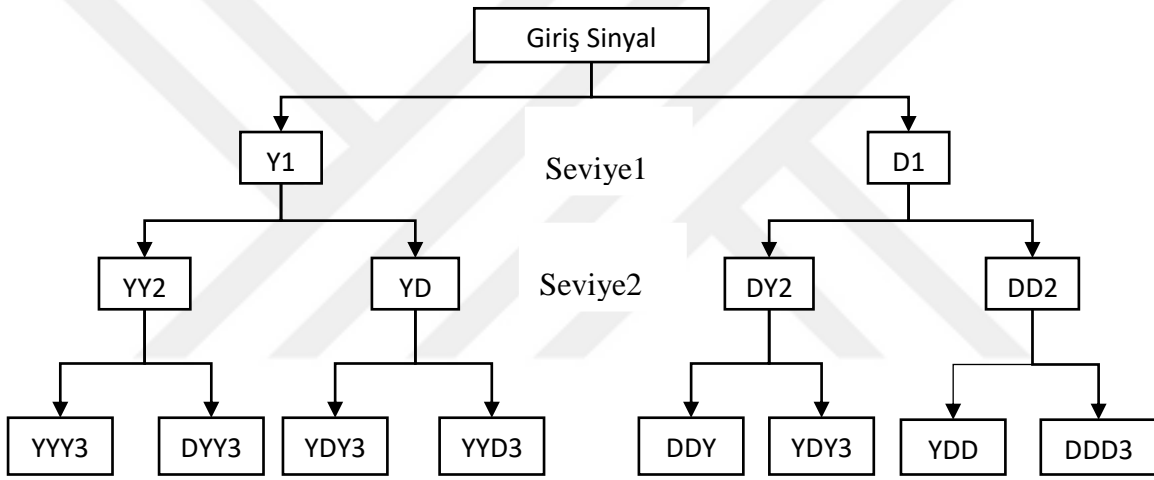
1.6.2. Frekans Domeninde Öznitelik Çıkarımı ve Analizi

EKG işaretleri frekans bileşenlerinin elde edilmesi için en yaygın tekniklerden biri zaman domeninde EKG işaretleri uygulanmış olan Hızlı Fourier dönüşümüdür (HFD). Fourier dönüşümüyle incelenen işaretler periyodik sinüzoid dalgalarına ayrıştırılır [139]. Böyle işaretler periyodik olduğu ve sonsuza kadar devam ettiği için, Fourier analizi periyodik veya istatistiksel karakteristikleri zamanla değişmeyen işaretler için uygundur. Fourier analizi frekans bileşenlerinin zamanla değiştiği, durağan olmayan işaretler için çok uygun değildir [140]. EKG işaretinin Fourier domain özellikleri, çoklu vuru tanıma [141], onset (dalga başlangıcı) tanıma [142], EKG 4-9 Hz frekans bandında güç spektrumu kullanarak atriyal fibrilasyon algılama [143,144], 4-7 Hz aralığı kullanılarak ventriküler fibrilasyon algılamada kullanılmıştır [145–147].

1.6.1.2 başlığında belirtildiği gibi EKG işaretleri için yaygın analiz frekans aralığı 100 Hz altındadır. Ancak son 40 yıl içerisinde EKG işaretlerine ait yüksek frekans bileşenleri de incelenmiştir [53]. EKG işaretlerinin yüksek frekans bileşenleri, hipertrofi, ventriküler anevrizması gibi, yukarıda belirtilen, yaygın EKG frekans aralığında belirgin olmayan anormallikleri algılamak için kullanılabilir [148].

1.6.3. Dalgacık Domeninde Öznitelik Çıkarımı ve Analizi

Fourier analizi ile ilgili çalışmalar on dokuzuncu yüzyılın başından beri yapılmışken, dalgacık dönüşümünün gelişimi çok daha yakın tarihte olmuştur. Haar dalgacığı olarak bilinen ilk dalgacık kavramı Alfred Haar tarafından 1909'da önerilmesine rağmen ilk uygulaması 1980'de gerçekleştirilmiştir. Dalgacık analizi iklim analizi, finansal analiz, işaret işleme, video görüntü işleme, veri sıkıştırma gibi birçok alanda yaygın olarak uygulanmıştır. Dalgacık dönüşümü Fourier dönüşümünden oldukça farklı benzersiz özelliklere sahiptir. Dalgacık dönüşümünün dalgacık olarak bilinen birçok temel fonksiyonu ve daha yerel özellikleri vardır [149].



Şekil 9. Dalgacık ayrıştırma ağacı. Y: İşaretin yaklaşım bilgisi D: İşaretin detay bilgisi

Dalgacık dönüşümü için dalgacık işaretleri kaydırma ve ölçekleme olarak iki şekilde düzenlenir. Kaydırma işleminde dalgacık, işaret boyunca çeşitli noktalara hareket ettirilir, ölçekleme işleminde ise dalgacık işareti genişletilir veya sıkıştırılır. Dalgacık dönüşümü dalgacık işareti ile dönüşümü yapılan işaretin yerel eşleşme miktarını hesaplar. Elde edilen sonuç kaydırma (x eksen) ve ölçek (y eksen) olarak iki boyutlu düzlemde ifade edilir. X eksen kaymayı gösterdiği için yüksek veya düşük eşleşmelerin tam olarak nerede olduğunu belirtir. Diğer bir deyişle Fourier dönüşümünden farklı olarak dalgacık dönüşümünde dönüşüm sonrası zaman bilgisine ulaşılabilir. Bu dalgacık dönüşümü için büyük bir avantajdır ve durağan olmayan işaretlerin analizi için de uygundur [150]. Dalgacık dönüşümünün diğer önemli bir özelliği ise çoklu çözünürlüktür. Dalgacık dönüşümü ile

zaman domenindeki işaretler alçak geçiren ve yüksek geçiren filtrelerle benzer olarak *yaklaşım* olarak adlandırılan düşük frekanslı bileşenlerine ve *detay* olarak adlandırılan yüksek frekanslı bileşenlerine ayrılmıştır. Bu bağlamda, dalgacık analizi frekans alanı analizine oldukça benzer [149,151]. Çoklu çözünürlüğün temelini oluşturan yaklaşım ve detay daha fazla ayrıştırılabilir.

Şekil 9’da giriş işaretini farklı seviyelerde ayrıştıran dalgacık ayrıştırma ağacı gösterilmiştir. EKG işareti dalgacık dönüşümü ile ayrıştırılabilir. Özellikleri çıkarılacak P ve R dalgası gibi güç spektrumuna göre tanıma için uygun ayrıştırma seviyesi seçilir [136]. EKG işaretleri için dalgacık domen analizi filtreleme ve sıkıştırma gibi işaret ön işleme ve öznitelik çıkarımı gibi işlem adımlarında kullanılmıştır [49,67,77,102,106,152,153].

Dalgacık ayrıştırma katsayıları analiz için doğrudan öznitelik olarak kullanılabilir. Örneğin Castro vd. ile Zhao vd. EKG işaretinin her bir vurusundan elde edilmiş dalgacık ayrıştırma katsayılarını öznitelik olarak kullanmışlardır [59,84]. Dalgacık ayrıştırma katsayıları yerel olarak zamanla ilişkilidir. Castro vd. P, QRS ve T dalgası ile ilgili yaklaşım katsayılarını çıkararak öznitelik vektörü olarak seçen bir algoritma önermiştir [59]. Zhao vd. bir kalp döngüsü ile ilgili aktiviteleri gösteren ayrıştırma katsayılarını öznitelik olarak kullanmıştır [84]. Detay ayrıştırma katsayıları EKG’nin Q, R ve S gibi dalgaları hakkında bilgi çıkarmak için kullanılmıştır [60]. Uygun dalgacık ve ayrıştırma seviyelerini bulmak dalgacıkların başarılı bir sonuç verebilmesi için önemlidir.

1.7. Boyut Azaltma

Öznitelik seçimi ve çıkarımı boyut azaltmada büyük önem arz etmektedir. Sınıflandırmada kullanılacak her bir öznitelik hesaplama maliyetini ve sistemin çalışma zamanını arttıracığı için daha az öznitelikle çalışmayı sağlayacak sistem ve modellerin geliştirilmesi önem taşımaktadır. Aynı zamanda seçilecek öznitelik alt kümesinin kabul edilebilir derecede yüksek performansa sahip sonuçlar üretmesi beklenir. Bu da en iyi öznitelik alt kümesini bulacak farklı arama tekniklerinin geliştirilmesine yol açmıştır. Öznitelik seçim yöntemlerinde özniteliklerin bir alt kümesinin seçimi ve boyut azaltmadan kaynaklanan bilgi kaybının en aza indirilmesine çalışılır. Çok sayıda öznitelik olduğunda tüm öznitelik alt küme ihtimallerinin denemesi hesaplama olarak kısıtlayıcıdır.

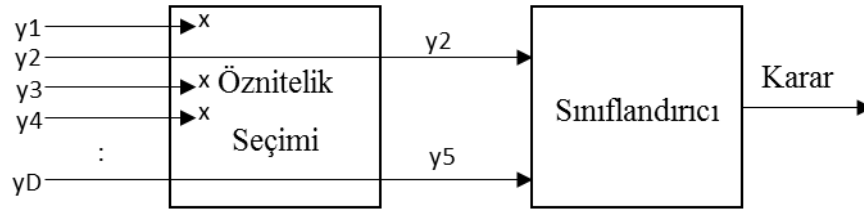
Örüntü tanımada öznitelik seçiminin üç ana amacı vardır: çıkarılan özniteliklerin hesaplama maliyetini azaltmak, sınıflandırma doğruluğunu arttırmak ve tahmin performansını geliştirmek [154]. Özellikle işaret işleme ve görüntü işleme ile ilgili örüntü tanıma problemlerinde başlangıçta seçilmiş aday öznitelikler büyük sayılarda olabilmektedir. Bu şekilde büyük sayılarda öznitelikler içlerinde değersiz öznitelikleri de barındırırlar. Bu öznitelikler sadece sınıflandırma için faydasız olmakla birlikte sınırlı sayıda örnekle eğitilmiş sınıflandırma şemalarının performansını da aşağıya çekmektedir. Böyle bir durumda gereksiz öznitelikler kaldırılarak sınıflandırma başarımı artırılabilir.

Öznitelik çıkarımı ise özgün özniteliklerin kombinasyon ve dönüşüm yöntemleri ile bazı yapay öznitelikler elde etmeye çalışır. Burada yeni hesaplanan öznitelikler önemli bilgileri özetleyen veya gizli bilgileri ortaya çıkaran özelliklerde olması istenir.

1.7.1. Öznitelik Seçimi

Öznitelik çıkarma işlemleri sonucunda işaretten elde edilen öznitelik kümesi ile sınıflandırılacak verinin boyutu önemli miktarda küçültülmektedir. Ancak elde edilen bu öznitelik kümesi içerisindeki bazı öznitelikler işaret ile ilişkisiz olabileceği gibi bazıları da işareti temsil etmede fazlalık oluşturabilir. Gereksiz olan ve fazlalık teşkil eden bu öznitelikler kullanılacak sınıflandırma yönteminin genelleme ve ayırıştırma yeteneğini olumsuz yönde etkileyebilir. Öznitelik seçme aşamasında öznitelik kümesinden bu tür istenmeyen özniteliklerin çeşitli yöntemler kullanılarak sadeleştirilip daha iyi bir öznitelik alt kümesi bulunması amaçlanmaktadır [155].

Öznitelik seçmenin ana amacı D boyutlu bir öznitelik uzayından d ($d < D$) boyutlu bir öznitelik uzayını, sınıflandırma sisteminin başarımını önemli miktarda azaltmadan veya arttıracak şekilde seçmektedir. Şekil 10'da öznitelik seçimini gösteren diyagram verilmiştir. Burada Y , D boyutlu bir öznitelik vektörünü, Y' ise d boyutlu bir öznitelik vektörünü göstermektedir. Örnek şekilde $d=2$ olarak seçilmiş ve bunlar giriş vektörünün 2. ve 5. elemanları olarak belirlenmiştir.



Şekil 10. Öznitelik Seçimi

Öznitelik alt uzayı seçme yöntemlerine göre öznitelik seçimi yöntemleri iki başlık altında toplanabilir.

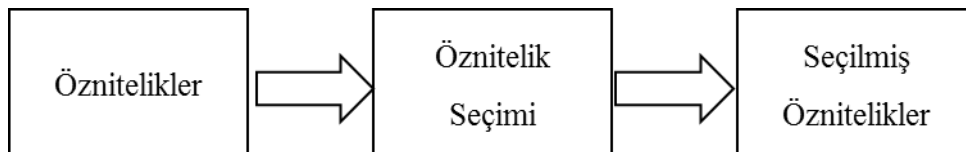
Optimum yöntemler: Bu yöntemler yorucu arama yöntemlerini içeren ve genellikle küçük boyutlu problemlerin çözümünde kullanılan yöntemlerdir. Çoğunlukla Somol vd. [156] geliştirdiği (Branch & Bound) şube&sınır prensibi üzerine kurulmuştur. Tüm optimum yöntemlerin yüksek boyutlu problemlerin çözümünde oldukça yavaş olmaları beklenir.

Yarı-Optimal yöntemler: Yarı optimum yöntemler hesaplama verimliliği için seçilecek olan alt öznitelik kümesinin optimum olmasına yoğunlaşırlar. Bunlara örnek olarak sıralı ileri ve sıralı geri arama, l-ekle-r-bırak (Plus-l-TakeAway-r), genetik algoritmalar ve bu algoritmaların geliştirilmiş sürümleri verilebilir [157].

Derinlemesine arama yapan algoritmalar çözümün optimum olmasına rağmen birçok gerçek hayat probleminde hesaplama yükü olarak kısıtlayıcıdır.

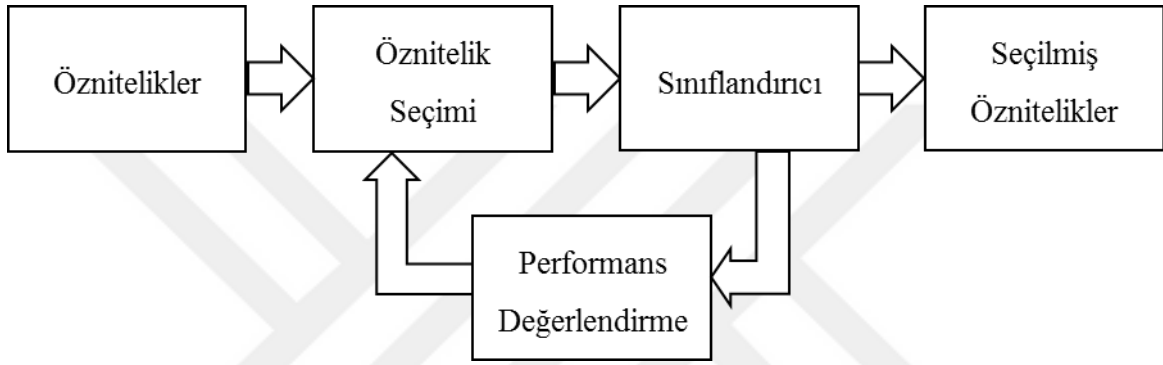
Öznitelikleri seçme kriterine göre ise öznitelik seçim yöntemleri kabaca üç başlık altında toplanabilir [157].

Filtreleme yöntemleri: Temel olarak eğitim verisi üzerinden doğrudan hesaplanan uzaklık, bilgi, bağımlılık, tutarlılık gibi performans değerlendirme fonksiyonları kullanılır ve alt öznitelik kümesini herhangi bir öğrenme algoritması kullanmadan seçilir. Filtre kullanan yöntemler hesaplama yükünü azaltmasına rağmen genellikle en iyi alt öznitelik kümesini bulamazlar [157].



Şekil 11. Filtre yöntemleri ile öznitelik seçimi

Paketleyici yöntemler: Bu yöntemlerde önceden belirlenmiş bir öğrenme algoritmasına ihtiyaç duyulur ve değerlendirme kriteri olarak bu algoritmanın performansı kullanılır. Genellikle paketleyici yöntemler filtre kullanan yöntemlerden daha iyi performans elde ederler. Ancak hesaplama yükü olarak çok daha maliyetlidir. Öznitelik alt kümesinin araştırılmasında sıralı ileri ve sıralı geri arama algoritmaları ve genetik algoritmalar gibi yöntemler yaygın olarak kullanılır [157]. Tez çalışmasında bu üç arama yöntemi karşılaştırmalı olarak kullanılmış ve sonuçları irdelenmiştir.



Şekil 12. Paketleyici yöntemler ile öznitelik seçimi

Gömülü Yöntemler: Bu yöntemler öznitelik seçimi işlemi ile model tahmini yöntemlerini birleştirirler. Böylece model tasarlama ve öznitelik seçimi birleşik bir öğrenme süreci oluşturur. Bu kategorideki yöntemler paketleyici yöntemlerin özel bir formu olarak düşünülebilir. Gömülü yöntemler hızlı öğrenme işlemleri sunarak paketleyici yöntemlere göre performans üstünlüğü sağlayabilirler. Ancak sonuçları üretmek için seçilmiş özel bir model ile uyum sağlaması gerekmektedir [157].

Sıralı Arama Algoritmaları ile Öznitelik Seçimi

Sıralı ileri arama (SİA) yöntemleri boş bir öznitelik kümesi ile başlar. Her bir adımda bazı ölçüt fonksiyonlarını sağlayan en iyi öznitelik kümeye dâhil edilir ve bu işlem n kez veya tüm öznitelikler için tekrarlanır. Böylece en yüksek ölçütü veren en iyi öznitelik alt kümesi seçilir. Sıralı geri arama (SGA) algoritması tüm öznitelikleri kullanarak aramaya başlar. Her adımda ölçüt fonksiyonuna göre en kötü öznitelik, öznitelik kümesinden çıkarılır

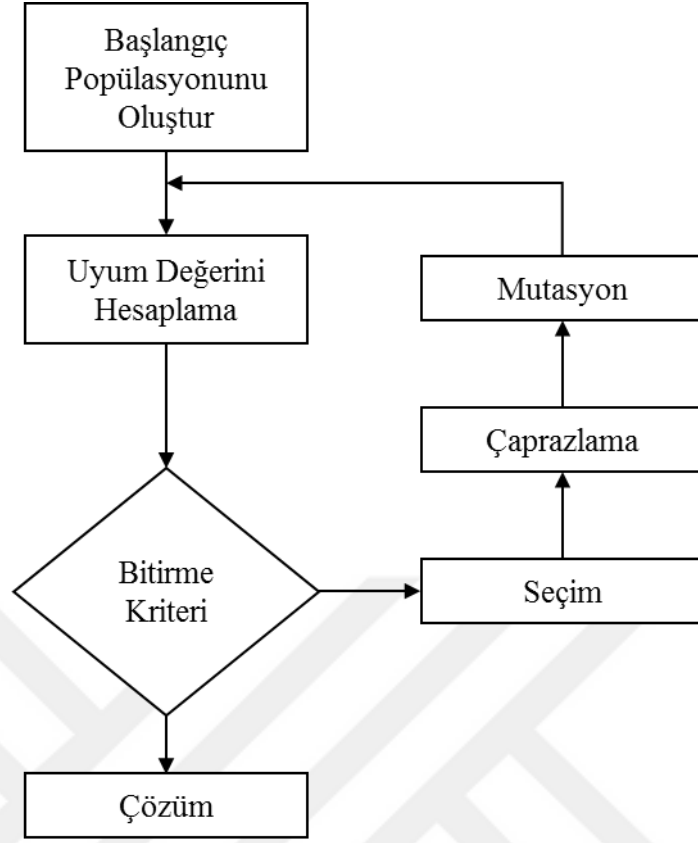
ve bu işlem r defa veya tüm öznitelikler için tekrarlanır. Böylece tüm öznitelikler içerisinde en iyi alt küme seçilmiş olur.

Sıralı arama yöntemlerinin genişletilmiş sürümü n -al r -gönder (n-take r -away veya Plus-1-Minus- r) olarak bilinir ve SİA ve SGA yöntemlerinin her ikisini de kullanır [158]. Bu algoritma boş bir öznitelik kümesi ile başlar. İleri arama durumunda her adımda bazı arama ölçüt işlevini sağlayan en iyi öznitelik geçerli öznitelik kümesine eklenir ve bu n kez tekrar edilir. Geri arama durumunda en kötü öznitelik elenir ve bu r kez tekrar edilir. Arzu edilen öznitelik boyutuna gelinceye veya ölçüt fonksiyonu düşmeye başlayıncaya kadar ileri arama dinamik olarak öznitelik sayısını arttırır ve geri arama öznitelik sayısını azaltır.

Genetik Algoritma ile Öznitelik Seçimi

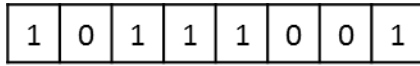
Genetik algoritmalar (GA) 1975 yılında John Holland tarafından geliştirilmiştir ve özel bir ortamdaki nüfusun evrimi için bir modeldir [159]. Nüfusun her bir üyesi gen dizisinden oluşan bir kromozom olarak gösterilir. Her bir genin iki veya daha fazla olası değeri vardır ve her gen problem uzayının bir parametresi haline dönüştürülmüştür. Bir uygunluk fonksiyonu ortamı temsil etmektedir ve her bireyi değerlendirerek uyum değerini tanımlar [38].

Algoritma popülasyon olarak adlandırılan bir dizi çözüm olarak sunulan kromozomlar ile çözüme başlar. Bir popülasyondan çözümler alınır ve yeni popülasyonlar oluşturmak için kullanılır. Yeni popülasyon oluşturma, seçme, çaprazlama ve mutasyon süreci uygulanır. Daha sonra uygunluk değeri hesaplanır. Bu işlemler en iyi sonuca ulaşma, belirlenen popülasyona ulaşma veya işlem süresi gibi bazı kriterlere ulaşıncaya kadar tekrarlanır. En azından bir en iyi sonuç değiştirilmeden yeni popülasyona kopyalandığı için yeni popülasyon eski olandan daha iyi olacaktır. Bu elitist stratejisi olarak bilinir. Şekil 13'te GA genel yapısı gösterilmiştir.

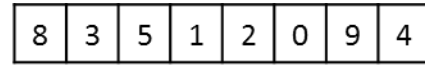


Şekil 13. GA genel yapısı

Kromozomlar ikili kodlama ve gerçel kodlama olarak iki yöntem ile kategorize edilebilir. Örneğin Şekil 14'deki gibi kaydedilen veri bit dizisi veya tam sayı dizisi olarak gösterilebilir. Veri ikili kodlamada 0 ve 1'lerden oluşur. İkili diziler parametre değerinde kullanılacakları zaman tam sayıya, ondalıklı sayıya veya GA uyum fonksiyonuyla kullanılacak herhangi bir parametreye dönüştürülür.



(a)



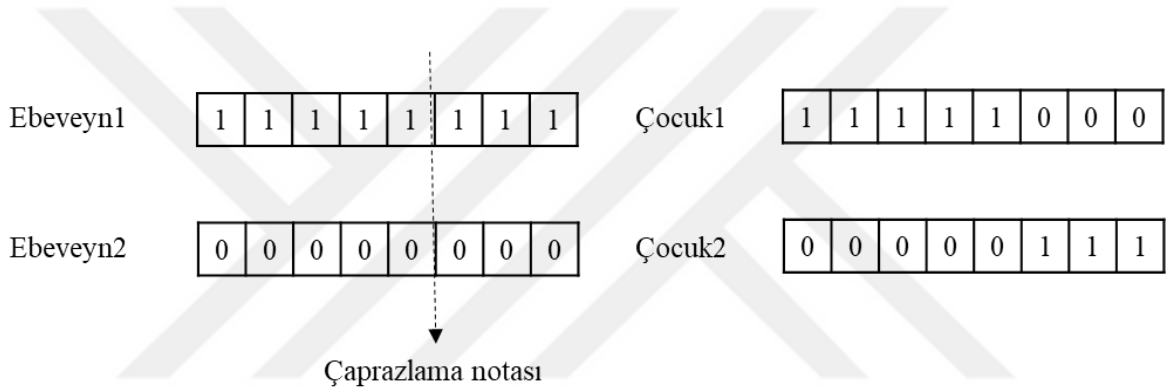
(b)

Şekil 14. Kromozomların gösterimi. (a) ikili kodlama, (b) gerçel kodlama

Her bir aday çözüm uyum fonksiyonuna göre optimizasyon problemi olarak değerlendirilir. Uyum fonksiyonu aday çözümün ne kadar iyi bir aday olduğunu tanımlayan bir ölçüt fonksiyonudur. GA popülasyon üyelerini daha iyi uyum fonksiyonu değerini verenleri bulmak için popülasyon içinde arama yapar.

Seçim önceki iki diziden yeni dizi oluşturmak için seçen bir operatördür. Yüksek uyum değerini veren bir dizinin seçilme ihtimali çok yüksektir. GA ebeveyn rasgele seçilir. Her bir dizinin uyum değeri seçilme olasılığını hesaplamak için kullanılır. Rulet tekerleği seçim şeması ve ranka dayalı seçim şeması sıklıkla kullanılır.

Çaprazlama ebeveyn dizilerden yavru dizeler oluşturmak için bir operatördür. Literatürde GA için çeşitli çaprazlama operatörleri önerilmiştir. İkili kodlama için Tek-Nokta çaprazlaması standart çaprazlama operatörlerine bir örnektir. Popülasyondan rasgele iki ebeveyn seçilir. Sonra bir çaprazlama noktası rasgele seçilir ve iki yeni dizi çaprazlama noktasından değiştirilerek iki yeni dizi oluşturulur. Tek nokta çaprazlama Şekil 15'te gösterilmektedir.



Şekil 15. İkili diziler için standart tek nokta çaprazlama

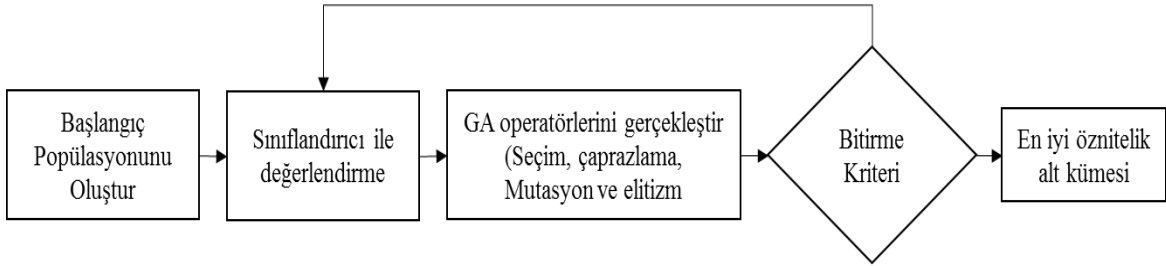
Mutasyon düşük olasılıklı bir çaprazlama operatörü tarafından oluşturulan dizinin elemanlarını değiştiren bir operatördür. GA'daki dizilere mutasyon operatörü uygularken dizideki rasgele seçilmiş bir bit değiştirilir. Şekil 16'daki mutasyon örneğinde görüldüğü gibi yıldızla gösterilen mutasyon noktalarındaki 1'ler 0'lar ile değiştirilmiştir. Mutasyon genetik çeşitlilik sağlar ve GA'nın daha geniş bir alanda araştırma yapmasına olanak sağlar.



Şekil 16. İkili dizi için iki bit mutasyon

Genetik algoritmanın öznitelik seçiminde uygulanması Şekil 17'deki akış diyagramında basit olarak gösterilmiştir. Burada popülasyondaki her bir öznitelik ikili bir

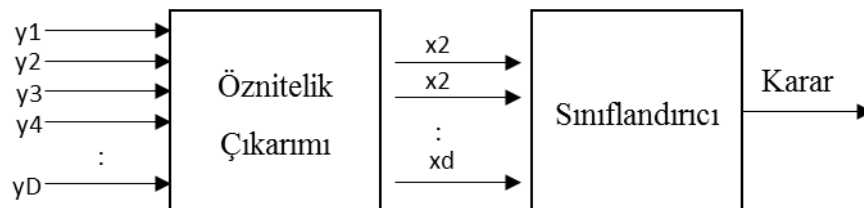
değer olarak gösterilmiştir. Değerlendirmeyi gerçekleştirecek olan uyum fonksiyonu sınıflandırma performansdır [24,27,38,160].



Şekil 17. GA ile öznelik seçimi

1.7.2. Öznelik Azaltma

Öznelik azaltma genellikle alt uzay projeksiyonu ile elde edilir. Birçok doğrusal alt uzay projeksiyon yöntemi ve bunların non-linear versiyonları vardır. Bu yöntemler özneliklerin birleşimini kullanarak azaltılmış boyutlu yeni öznelikler oluşturulmasını sağlar ve özneliklerin fizyolojik olarak anlamları korunmaz. Temel bileşen analizi (TBA) boyut indirgemede yaygın olarak kullanılmaktadır. Ancak bulunan projeksiyonlar varyansı maksimize ederler, sınıflandırma başarımı ile doğrudan ilişkili olmak zorunda değildir. Doğrusal diskriminant analizi (DDA) Gauss dağılımı varsayımı altında sınıfların ayrılmasını maksimize eden lineer projeksiyonları bularak TBA'nın bu eksikliğini gidermeye çalışır [161]. DDA projeksiyonları sınıflara ait ortalama ve kovaryans matrislerine dayanarak optimize edilmiştir. Bağımsız bileşen analizi (BBA) bağımsız değişkenlerin istatistiksel olarak bağımsızlıklarını maksimize eden lineer dönüşümleri bulan bir araç olarak kullanılmaktadır [162]. Şekil 18'de öznelik çıkarımı gösterilmektedir. Burada giriş vektöründen farklı bir yöntem kullanılarak elde edilmiş ve boyut olarak daha küçük değere sahip bir vektör bulunmaktadır.



Şekil 18. Öznelik çıkarımı

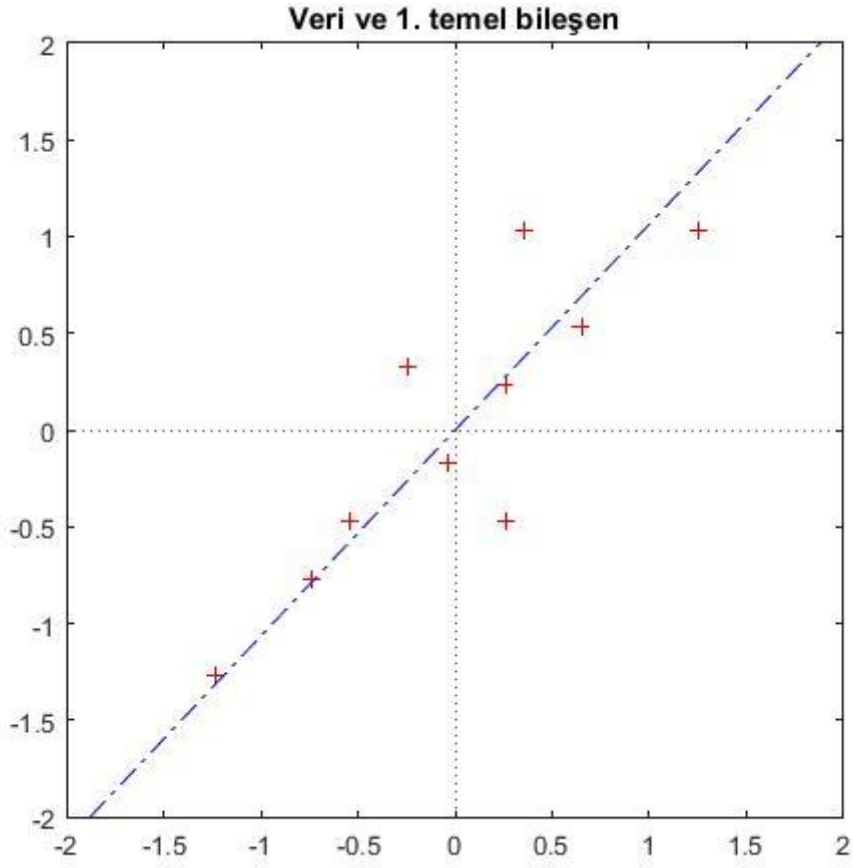
Büyük veri setlerinde boyut indirgeme için en popüler yöntem, Karhunen-Loeve (K-L) olarak bilinen temel bileşen analizidir (TBA). Çeşitli alanlarda, tekil değer ayrışımı (TDA), Hotelling dönüşümü ve ampirik ortogonal fonksiyonu olarak bilinir. TBA başlangıçta verilmiş vektör örneklerini, boyutu indirgenmiş yeni vektör örneklerine dönüştüren bir dönüşüm yöntemidir. Bu dönüşümün amacı boyutu düşürülmüş veri ile gerçek örnekler arasındaki bilgi değişimine yoğunlaşmaktır. Pratik uygulamalar TBA'nin ortalama karesel hata (OKH) anlamında en iyi lineer boyut azaltma tekniği olduğunu doğrulamaktadır [163]. Özniteliklerin kovaryans matrisine dayalı olarak ikinci mertebeden bir yöntemdir. Esasında TBA, en büyük varyans ile orijinal özniteliklere birkaç doğrusal kombinasyonlarını bularak verilerin boyutunu azaltmayı amaçlamaktadır. Varyans değişkenlerin ölçeğine bağlı olduğundan her bir değişken öncelikle 0 ortalamalı ve 1 varyanslı olacak şekilde standartlaştırılır.

Temel Bileşen Analizi

TBA şu şekilde ifade edilebilir: $X = \{x_1, x_2, x_3, \dots, x_m\}$ m-boyutlu bir vektör $Y = \{y_1, y_2, y_3, \dots, y_m\}$ aynı boyuta sahip farklı bir vektöre dönüştürülebilir. Ama burada y'ler bilgi içeriğinin çoğunu ilk birkaç ögesinde tutarlar. Boyutu en az bilgi kaybı olacak şekilde azaltmak için bu bilgiden faydalanılır ve en az kayıpla vektörün boyutu azaltılmış olur. Dönüşüm yüksek varyansın yüksek bilgiye karşılık geldiği varsayımına dayanmaktadır. Bu yüzden X vektörünün boyutu azaltılmak isteniyorsa X 'i tek boyutlu Y 'ye dönüştürecek aşağıdaki dönüşüm işlemini gerçekleştirmelidir.

$$Y=A*X \quad (2)$$

Burada A verilen veri için mümkün olan en yüksek varyansı elde edecek şekilde seçilmelidir. Dönüşümden elde edilen tek boyutlu Y ilk temel bileşen olarak adlandırılır. İlk temel bileşen en yüksek varyansın yönünde bir eksenidir. Bu eksen veri noktaları ile ilk temel bileşen eksenine izdüşümlerinin karelerinin toplamı mesafesini en küçük değerine indirir [164].



Şekil 19. Örnek veri ve hesaplanmış 1. Temel bileşenin yönü

Şekil 19’da görüldüğü gibi iki boyutlu uzay veri setinin en yüksek varyansına sahip tek boyutlu uzaya dönüştürülmüştür. Uygulamada A matrisini doğrudan tanımlamak mümkün değildir. Bu nedenle ilk aşamada özneliklere ait kovaryans matrisi S hesaplanmaktadır. Kovaryans matrisini hesaplamak için (3) eşitliği kullanılır.

$$S_{n \times n} = \frac{\sum_{j=1}^n (x_j - \bar{x})^T (x_j - \bar{x})}{n - 1} \quad (3)$$

Burada \bar{x} (4) eşitliği ile hesaplanan ortalama vektörüdür.

$$\bar{x} = \frac{\sum_{j=1}^n x_j}{n} \quad (4)$$

Bir sonraki adımda verilen verinin kovaryans matrisinin özdeğerleri hesaplanır. Sonuç olarak kovaryans matrisindeki m adet en geniş özdeğerlere karşılık m adet özvektör hesaplanır. Böylece n -boyutlu uzaydan m -boyutlu uzaya öznitelikler ilişkisiz olacak şekilde dönüşüm tamamlanmış olur. Bağımsız bileşenlerin tanımlamasını tamamlamak için şu notasyonları da belirtmek gerekir.

S matrisinin özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_n$ olarak gösterilir ve $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ koşulunu sağlamaktadır. Özvektörler e_1, e_2, \dots, e_n olarak gösterilir ve temel eksenler olarak adlandırılırlar. Temel eksenler yeni bir n -boyutlu uzaya dönüştürülmüş eksenlerdir. Burada yeni değişkenler ilişkisizdir ve i . bileşenin varyansı i . özdeğere eşittir. Çünkü λ_i 'ler sıralıdır ve veri setinin içerdiği bilginin çoğu ilk birkaç temel bileşende yoğunlaşmıştır. Burada önemli soru kaç tane temel bileşenin veriyi en iyi temsil edebileceğidir. Diğer bir ifadeyle veri setinin en etkili boyutu kaçtır? Bu soruyu cevaplamanın en kolay yolu varyans oranını analiz etmektir. İlk m tane özdeğerin toplamı tüm özdeğerlerin toplamına bölüldüğünde ilk m temel bileşenin temsil etme kalitesi ölçülmüş olacaktır. Sonuç % olarak hesaplanır ve örnek olarak verilecek olursa %90 ve üzeri çıkıyorsa hesaplanan değer temsil yeteneği iyi denebilir. Bu işlem aşağıdaki (5) eşitliği ile hesaplanır.

$$R = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (5)$$

Hesaplanan R değeri belirlenen bir eşik değerinden yüksek olduğunda özniteliklerin m adet alt kümesi b -boyutlu uzay için iyi bir tahmini temsil eder. Bu yöntem hesaplama yükü olarak ucuzdur. Ancak verinin kovaryans matrisi olarak hesaplanmasını gerektirir.

Bağımsız Bileşen Analizi

TBA uygulanarak birbiri ile ilişkisiz öznitelikleri oluşturmaktadır. Amaç boyut indirgeme ve ortalama kare hatanın minimize edilmesi ise TBA ile bulunan çözüm uygundur. TBA'den daha sonra geliştirilmiş bağımsız bileşen analizi (BBA) veriyi ilişkisiz (dekorelasyon) duruma getirmeden biraz daha fazlasını elde etmeye çalışmaktadır [165]. Çok boyutlu istatistiksel verilerde gizli bileşenlerin bulunması için istatistiksel bir modeldir. BBA doğrusal olmayan bir boyut indirgeme yöntemidir [166]. BBA modelinde gizli

bileşenler bilinmeyen bir karıştırma mekanizmasıyla karıştırılarak gözlem verileri oluşturulduğu varsayılır. BBA'da temel doğrusal karışım modeli eşitlik (6)'daki gibi ifade edilir.

$$Y=W \cdot X \quad (6)$$

Giriş örnekleri X kümesi göz önüne alındığında, $Y = \{y_1, y_2, y_3, \dots, y_n\}$ dönüştürülmüş vektörü karşılıklı olarak bağımsız yapan $n \times n$ boyutlu bir W tanımlanabilir. İstatistiksel bağımsızlık hedefi TBA tarafından sağlanan ilişkisizlikten daha güçlü bir durumdur [164]. Bu iki koşul sadece Gauss rasgele değişkenler için eşdeğerdir.

İlişkisizlikten ziyade bağımsız özniteliklerin aranıyor olması verinin yüksek mertebeden istatistiklerinde gizlenmiş daha fazla bilginin ortaya çıkmasını sağlayacaktır. Giriş veri vektörümüz X istatistiksel olarak bağımsız ve durağan bileşenlerin kombinasyonu ile oluşturulduğunu varsayalım. Bu eşitlik (7)'de gösterilmiştir.

$$X=A \cdot Y \quad (7)$$

Artık işlem X 'te gizlenmiş bilgilerden yararlanılarak eşitlik (6)'daki Y bileşenlerini geri kazanmak için W matrisi hesaplanabilir. Burada A genellikle karışım matrisi, W 'de ayırışım matrisi olarak bilinir ve birbirlerinin tersleridir [166].

BBA modeli tanımlanabilirlik şartları vardır. Tüm bağımsız bileşenler $Y = \{y_1, y_2, y_3, \dots, y_n\}$ Gauss olmamalıdır. Bu birinci istisnadır. İkinci koşul ise A matrisi tersi alınabilir bir matris olmalıdır. A $L \times N$ gibi kare olmayan bir matris olması gibi daha genel bir durumda L , N 'den daha büyük olmalı ve A matrisinin rankı sütun sayısına eşit olmalıdır.

Diğer bir ifadeyle her durumda uygulanabilen TBA'den farklı olarak BBA yalnızca değişkenlerin Gauss olmadığı durumda anlamlıdır. Daha önce de belirtildiği gibi Gauss dağılımlı rasgele değişkenler için ilişkisizlik bağımsızlığa eşdeğerdir ve bu durumda TBA yeterlidir.

Çalışmada bağımsız bileşenlerin hesaplanması için hızlı BBA [167] algoritması kullanılmıştır. Aşağıda hesaplama adımları verilmiştir.

İlk aşama ön işleme aşamasıdır. Bu aşamada verinin ortalaması veriden çıkarılır ve böylece verinin ortalaması sıfırlanmış olur. Bu işleme merkezileştirme denir. Bu durumda bağımsız bileşenlerin ortalaması da sıfır olacaktır. Çıkarılmış bu ortalama değer karışım matrisi kestirim yapıldıktan sonra tekrar sonuca eklenerek kestirim işlemi tamamlanır.

$$X = X - E\{X\} = X - \frac{1}{N} \sum_{i=1}^N x_i \quad (8)$$

Burada $E\{.\}$ istatistiksel beklenen değerdir ve N ise veride sunulan örnek sayısıdır. BBA'de kullanılan diğer bir ön işlem beyazlatmadır. Beyazlatmanın amacı X veri matrisini ilişkisiz ve 1 varyanslı yeni bir \tilde{X} matrisine dönüştürmektir. \tilde{X} matrisi kovaryansı birim matrise eşit olan bir matristir. En yaygın kullanılan beyazlatma yöntemi X 'in kovaryans matrisinin özdeğer ayrıştırmasıdır.

$$\tilde{X} = VD^{-1}V^T X \quad (9)$$

Eşitlik 10'da eşitliğin sol tarafı ($E\{XX^T\}$) verinin kovaryans matrisidir, V ; $E\{XX^T\}$ 'nin özvektörlerinden oluşan ortogonal bir matristir ve D ise özdeğerlerden oluşan köşegen bir matristir. Beyazlatma işlemi (9) eşitliği ile gerçekleştirilir.

$$E\{XX^T\} = VDV^T \quad (10)$$

Böylece ön işleme adımı tamamlanmış olur. Hızlı BBA algoritmasında negentropi yaklaşımı kullanılır. Burada $W^T X$ Gauss-dışılığını maksimum yapacak bir W vektörü belirlenmeye çalışılır. Burada Gauss-dışılığı hesaplamak için negentropinin $J(W^T X)$ yaklaşımı kullanılır. Bu durumda ön işlemden geçmiş veriler için W 'nin normunun 1'e eşitlenmesi gerekmektedir.

Hızlı BBA Gauss-dışılığın en yüksek (maksimum) noktasını bulan iteratif yöntemeye dayanır. Bu aşamadan sonra hızlı BBA algoritmasının genel akışı aşağıda verilmiştir.

- Başlangıç ağırlık vektörü W 'yi seç (rasgele olabilir).
- $W^+ = E\{Xg(W^T X)\} - E\{g'(W^T X)\} W$ hesapla.

Burada $g(u) = \frac{1}{a_1} \log \cosh(a_1 u)$ ve $g'(u)$ g 'nin türevidir.

- W^+ 'yi aşağıdaki şekilde normalize et.

$$W = \frac{W^+}{\|W^+\|}$$

- Yakınsamadıysa yani bir önceki iterasyonda değeri değiştiyse 2. Adıma dön.

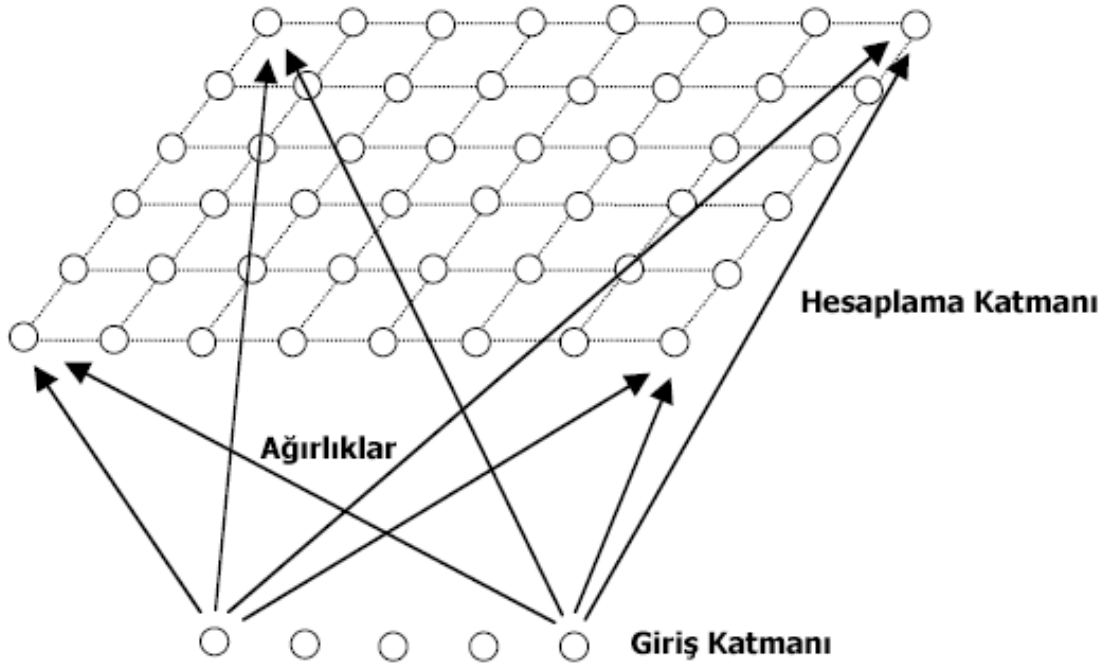
Yukarıdaki yöntemle W bulunduktan sonra tersi alınır ve A matrisi bulunmuş olur [166].

Özdüzenleyici Haritalar

Özdüzenleyici haritalar (ÖH) (Self-Organizing Map) veya Özdüzenleyici öznitelik haritaları (ÖÖH) giriş uzayındaki eğitim örneklerinin ayrıştırılmış sunumu olan düşük boyutlu çıkışı üretmek için danışmansız öğrenme yöntemini kullanan bir yapay sinir ağı türüdür. ÖH giriş uzayının topolojik özelliklerini korumak için bir komşuluk fonksiyonu kullandığı için diğer yapay sinir ağlarından ayrılır. Geometrik yapıda düzenlenmiş nöronların komşuluk ilişkisi, bu nöronları tanımlayan ağırlık vektörlerinin güncellenmesi sırasında dikkate alındığından, yakın komşuluktaki nöronlar benzer verileri temsil etmektedir. Böylece verilerin oluşturduğu giriş uzayının topolojik özellikleri bir veya iki boyutlu geometrik yapıya aktarılmış olur. Bu özelliği çok boyutlu ölçeklemeye benzer bir şekilde çok boyutlu verilerin görünümünü görselleştirmek için ÖH'ı faydalı yapmaktadır. Model ilk olarak Teuvo Kohonen tarafından yapay sinir ağı olarak tanımlanmıştır [168,169].

Yapay sinir ağlarının çoğu gibi ÖH iki modda çalışabilir: Eğitim ve haritalama. Eğitim giriş örneklerini kullanarak harita oluşturur. Haritalama otomatik olarak giriş vektörünü sınıflandırır. ÖH düğüm, nöron veya ünite olarak adlandırılan bileşenlerden oluşur. Her bir düğüm ilişkili olarak giriş vektörü ile aynı boyutta ağırlık vektörüne ve harita uzayında bir pozisyona sahiptir. Genellikle düğümlerin yerleşimi düzgün aralıklı altıgen veya dikdörtgen biçimli ızgara şeklindedir.

Eğitim aşamasında giriş verisinin ağa sunulması ile nöronların ağırlıkları ayarlanarak verilerin oluşturduğu grupları temsil edecek şekilde değişir. Yüksek boyutlu giriş vektöründeki her bir örnek kazanan birimle ilişkilendirilir. Sadece kazanan nöron değil komşuları da giriş değerine göre ağırlık değerlerini adapte ederek öğrenir. Şekil 20'de ÖH genel yapısı gösterilmiştir.



Şekil 20. ÖH Genel yapısı

Özyineleyici ağırların öğrenme algoritması aşağıdaki adımlardan oluşur.

İlk Değer Atama: Ağırlık vektörleri $w_i, i = 1, 2, \dots, n$ ilk değerleri rasgele atanır. Böylece verilerin küme merkezlerine denk gelen bir nöronun ilk değerleri atanmış olur. Bunlar genellikle standardize edilmiş küçük rasgele değerlerdir.

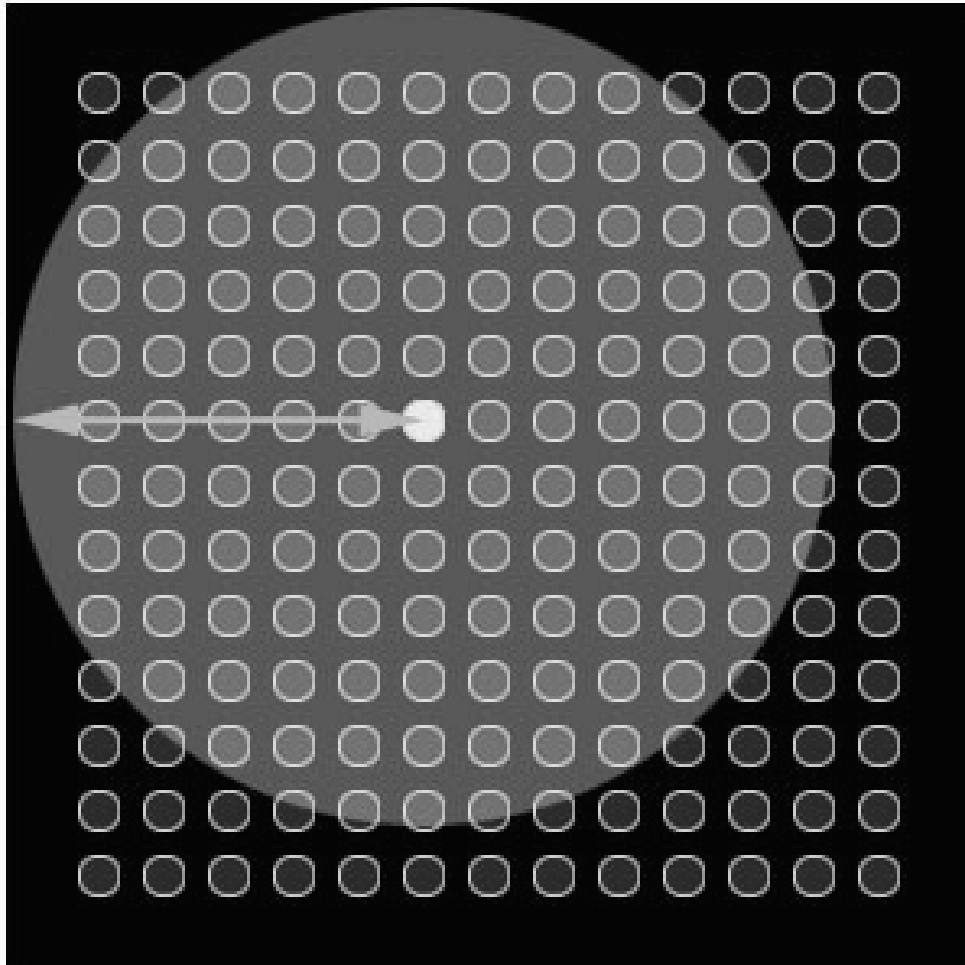
Kazanan birim hesaplama: Eğitim verisinden rasgele bir vektör seçilir ve ızgarada gösterilir. Her bir düğüm hangisinin ağırlığı giriş vektörünün değerine en yakın olduğunu hesaplamak için incelenir. Kazanan düğüm literatürde kazanan birim (BMU- Best Matching Unit) olarak bilinir. Kazanan birimi (KB) belirlemek için yöntemlerden bir tanesi, tüm düğümler üzerinde o anki düğümün ağırlık vektörü ile şu anki giriş vektörünün arasındaki öklit mesafesini hesaplamaktır. Giriş vektörüne en yakın ağırlık vektörüne sahip düğüm KB olarak belirlenir.

$$D = \sqrt{\sum_{i=1}^n (V_i - w_i)^2} \quad (11)$$

burada V şu anki giriş vektörü ve w düğümün ağırlık vektörüdür.

Komşuluk Hesapla: KB'in komşuluk yarıçapı hesaplanır. Komşuluk yarıçapı başlangıçta büyük olarak (ızgara yarıçapı kadar) alınır, ancak her bir adımda azaltılır. Bu yarıçap içerisinde bulunan düğümler KB'in komşusu sayılır. KB ağırlığı ve komşuluk yarıçapında bulunan komşuların ağırlıkları belirlenmiş komşuluk fonksiyonuna göre güncellenir. KB'e yakın olan komşular ağırlıktan daha çok etkilenirler.

Her döngüde, KB belirlendikten sonra bir sonraki adım diğer düğümlerden hangilerinin KB'in komşusu olacağı hesaplanması aşamasıdır. Belirlenen tüm bu düğümlerin ağırlık vektörleri değiştirilir. Burada hesaplanacak ilk değer komşuluğun yarıçapının ne olacağıdır. Şekil 21'de ilk başlangıçtaki kazanan birimin komşuları gösterilmektedir.

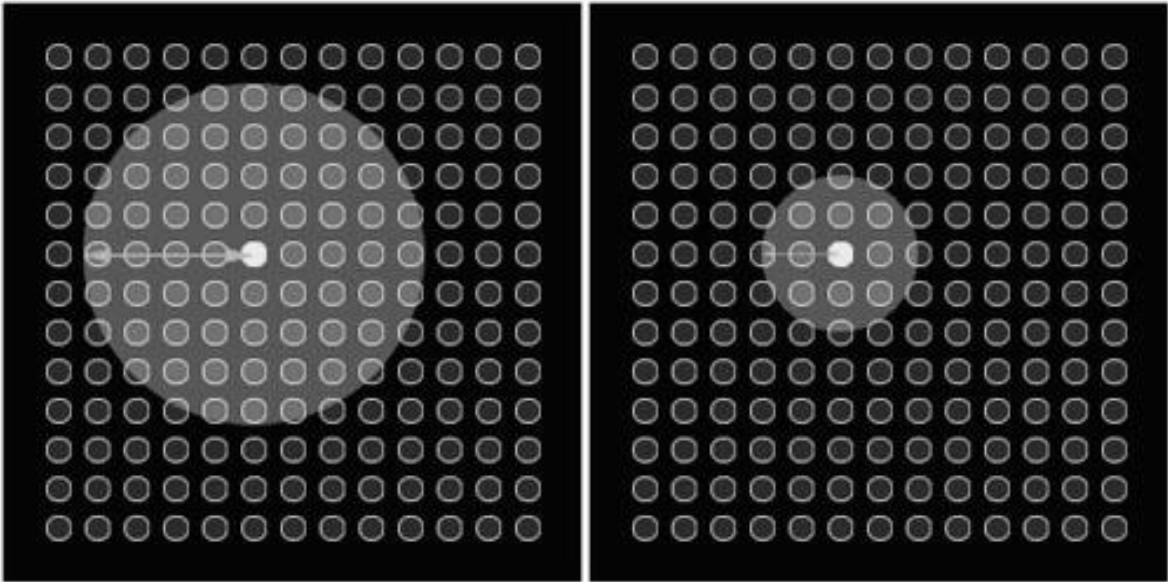


Şekil 21. Kazanan birimin komşulukları

Açık tondaki nokta KB'dir ve komşusu olarak belirlenecek birçok düğümün ağırlıklarını etkileyecektir. Ok ile komşuluk sınırını belirleyecek olan dairenin yarıçapı gösterilmektedir. Kohonen'in öğrenme algoritmasının benzersiz özelliklerinden bir tanesi komşuluk alanının zamanla küçülüyor olmasıdır. Bu işlem zamanla komşuluk yarıçapının daraltılmasıyla sağlanır. Komşuluk yarıçapı eşitlik (12) ile hesaplanır:

$$\sigma(t) = \sigma_0 e^{-\frac{t}{\lambda}}, \quad t = 1, 2, 3, \dots \quad (12)$$

Burada σ_0 , t_0 anındaki ızgara genişliğini göstermektedir. λ , zaman sabitini göstermektedir. t , aktif işlem adımını (döngüdeki iterasyon) göstermektedir.



Şekil 22. Komşuluk çemberinin daraltılması

Şekil 22'de komşuluk yarıçapının zamanla daraltılması gösterilmektedir. Zamanla komşuluk tek bir düğüm (KB) kadar azalacaktır. Açı hesaplandıktan sonra bir sonraki adıma geçilir. KB'in (KB dahil) tüm komşularının ağırlık vektörleri eşitlik (13)'e göre ayarlanır.

$$w(t+1) = w(t) + L(t)(V(t) - w(t)) \quad (13)$$

Burada t , aktif işlem adımını, L öğrenme oranı olarak adlandırılan ve zamanla değeri azalan öğrenme değişkenini, V şu anki giriş vektörünü göstermektedir. Öğrenme oranı L her bir adımda eşitlik (14) kullanılarak tekrar hesaplanır.

$$L(t) = L_0 e^{-\frac{t}{\lambda}}, \quad t = 1, 2, 3, \dots \quad (14)$$

Dikkat edilirse bu fonksiyon öğrenme katsayısının kullanılması haricinde eşitlik (12)'de verilen üstel fonksiyon ile aynıdır. L_0 başlangıçta belirlenen bir sabittir. Öğrenme katsayısı, hesaplamanın son adımlarında sıfıra yakın bir değere ulaşır.

(13) eşitliğinde sadece zamana göre bir azalma vardır. Bu formül sadece KB için doğru çalışacaktır. Kazanan düğümden uzaklaştıkça etkinin azalması ve kenara varıldığında en aza inmesi için formüle ufak bir ekleme yapmak gerekmektedir.

$$w(t+1) = w(t) + \Theta(t)L(t)(V(t) - w(t)) \quad (15)$$

Burada Θ düğümün kazanan düğümden olan uzaklığının öğrenmeye olan etkisini göstermektedir ve eşitlik (16)'da verilmiştir.

$$\Theta(t) = e^{-\frac{D^2}{2\sigma^2(t)}}, \quad t = 1, 2, 3, \dots \quad (16)$$

Burada D hesaplama yapılan düğümün kazanan düğüme olan uzaklığı, σ eşitlik (12)'de verilen komşuluk genişliğidir.

İterasyon Tekrar: İkinci adımdan itibaren iterasyon tekrarlanır. Eğitim kümesindeki aynı veriler için aynı nöronlar kazanmaya başladığında gruplaşma tamamlanmıştır. Grupların verileri temsil etme özelliklerinin etkinleşmesi için işlemlere belirlenen bir tekrar sayısı kadar devam edilir.

1.8. EKG Verilerinin Sınıflandırılması

Örüntü tanımada son aşama örüntünün sınıflandırılmasıdır. Örüntüyü bir kategoriye atamak için öznitelik çıkarma/azaltma aşamasında elde edilen öznitelik vektörünü kullanan

ve tüm sistemin başarılı çalışmasında rol oynayan en önemli bileşenidir [170]. Bilgisayar hesaplama güçlerindeki artış büyük verilerin sınıflandırılması için çok çeşitli yöntemlerin kullanılmasına olanak tanımıştır.

Örüntü sınıflandırmada öğrenme algoritmaları danışmanlı öğrenme veya danışmansız öğrenme olarak iki ayrı grupta toplanabilir. Danışmanlı öğrenmede bir danışman eğitim setindeki her örüntü için sınıf etiketi sağlar ve bu örüntüler için hatayı azaltacak değerleri arar. Diğer yandan danışmansız öğrenmede sınıf etiketleri kullanılmaz, örüntüler üzerindeki benzerlikleri kullanarak öğrenme işlemini gerçekleştirirler [170]. Bu bölümde, çalışmada kullanılan öğrenme algoritmalarından kısaca bahsedilecektir.

1.8.1. K-En Yakın Komşu Algoritması

K-En yakın komşu (K-NN) örüntü tanımada en etkili parametrik olmayan yöntemlerin başında gelir [164,170,171]. K-NN yöntemi en basit makine öğrenmesi algoritmasından biridir. Algoritma öğrenilen örneklerin istatistiksel dağılımına bağımlı değildir. K-NN algoritması öznitelik uzayında eğitilmiş örneklerden en yakınına bakarak sınıflandırma yapan bir yöntemdir. Komşuluk tanımlamasından dolayı çok boyutlu uzayda nesnelere konum vektörleri ile temsil edilirler. K-NN yönteminde sınıflandırılmak için yeni bir x örneği geldiğinde, en yakın k komşusu arasından sayıca en fazla olan sınıfa atanır. Algoritma farklı uzaklık ölçüm yöntemleri kullanmaktadır. *Öklid* uzaklığı yaygın olarak uzaklık ölçümü olarak kullanılmaktadır. $X = \{x_1, x_2, x_3, \dots, x_m\}$ ve $Y = \{y_1, y_2, y_3, \dots, y_n\}$ arasındaki uzaklık eşitlik (17) gösterilmiştir.

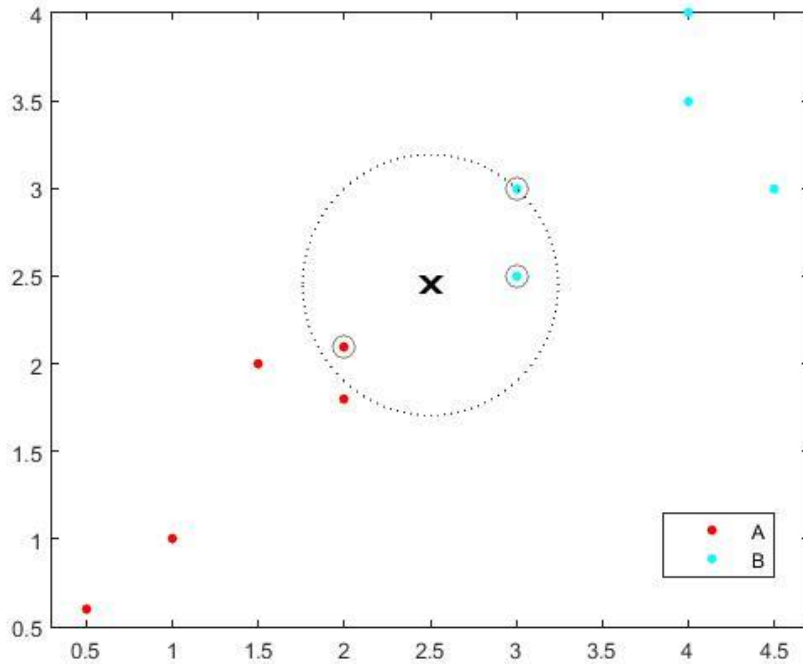
$$u = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (17)$$

Öklid uzaklığından farklı olarak *Manhattan*, *Minkowski* gibi farklı uzaklık ölçütlerini de kullanmak mümkündür. *Manhattan* uzaklığı eşitlik (18)'de, *Minkowski* uzaklığı ise eşitlik (19)'da gösterilmektedir.

$$u = \sqrt{\sum_{i=1}^n |x_i - y_i|} \quad (18)$$

$$u = \sqrt[q]{\sum_{i=1}^n (x_i - y_i)^q} \quad (19)$$

k değeri genellikle küçük değerler seçilmektedir. $k=1$ olarak seçildiğinde örnek noktaya en yakın komşusunun sınıfı sınıflandırma sonucu olarak atanacaktır. Burada dikkat edilmesi gereken k sayısına bağlı olarak sınıflandırmanın farklı sonuçları olabileceğidir. Şekil 23'te örnek bir K-NN sınıflandırma işlemi gösterilmiştir. Şekilde kırmızı (A sınıfı) ve turkuaz (B Sınıfı) renkli örnekler olmak üzere iki sınıf vardır. Test için gelen örnek ise x olarak işaretlenmiştir. Kesikli çizgilerle çizilmiş daire $k=3$ için çizilmiş kapsama alanıdır. Şekilden anlaşılacağı gibi test örneği iki tane turkuaz nokta olduğu için B sınıftan olacaktır. $k=5$ olarak belirlendiğinde ise soldaki iki kırmızı nokta daireye girecek ve sınıflandırma sonucu A sınıfı olarak değişecektir.



Şekil 23. K-NN sınıflandırma örneği

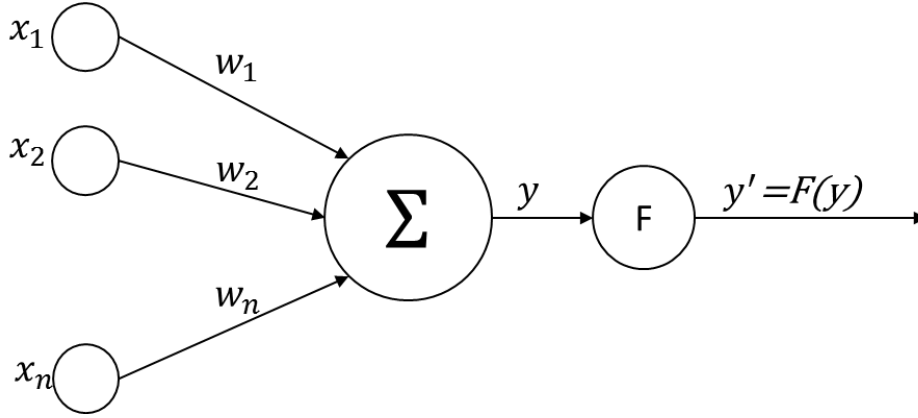
K-NN danışmanlı öğrenme yapan bir sınıflandırma algoritmasıdır. Yeni gelen örneğin sınıflandırılabilmesi için sınıf etiketine sahip örneklere ihtiyaç duymaktadır. Bu örnekler algoritmanın eğitim seti olarak düşünülebilir. Bunun haricinde herhangi bir ek eğitim işlemine gerek duyulmamaktadır.

K-NN algoritması kolay anlaşılabilir bir algoritma olduğu için herhangi bir probleme uyarlanması basittir. Sayısal veriler içeren bir veri seti üzerinde uygulanması kategorik veri içeren veri setlerine oranla daha kolaydır. Ayrıca gürültüye sahip veriler için de olumlu sonuçlar üretebilir. Veri boyutu arttığında algoritmada kullanılacak verilerin saklanması hafızada fazla yer kaplar. K-NN'in bir diğer dezavantajı yeni bir örnek sınıflandırırken eğitim setindeki diğer örneklere olan uzaklığının hesaplanması ve karşılaştırılmasının zaman almasıdır.

1.8.2. Yapay Sinir Ağları

Donald Hebb 1949 yılında yazdığı “The Organization of Behavior” isimli kitabında ortaya attığı fikirle Yapay Sinir Ağları (YSA)'nın fikir babası sayılır. Nörolog olan Hebb, beynin nasıl öğrendiği ile ilgili çalışmalar yapmıştır. YSA ile basit biyolojik sinir sisteminin çalışma biçimi taklit edilmiştir. Esinlenen sinir hücreleri nöronları içerir ve bu nöronlar farklı şekillerde birbirine bağlanarak sinir ağını oluştururlar. Yapay nöronlar ve yapay sinir ağları açısından bakıldığında, Hebb ilkesi model nöronları arasındaki ağırlıkların nasıl değiştirileceğini tanımlayan bir yöntem olarak açıklanabilir [172,173].

Yapay sinir hücresi (nöron), gerçek biyolojik sinir hücresi ile aynı ilkelere dayandırılmaya çalışılmıştır. YSA'lar, ağırlıklandırılmış şekilde birbirlerine bağlanmış birçok nörondan oluşan matematiksel yapılardır. Bir nöron, diğer nöronlardan girişleri alır; bunları birleştirir, dönüştürür ve sayısal bir sonuç ortaya çıkartır. Şekil 24'te bir yapay sinir hücresinin yapısı gösterilmiştir.



Şekil 24. Yapay sinir hücresinin yapısı

Şekil 24'te gösterilen yapay sinir hücresinin matematiksel ifadesi eşitlik (20)'de verilmiştir.

$$y = \sum_{i=1}^n w_i \cdot x_i \quad y' = F(y) = \begin{cases} 1 & y \geq 0 \\ 0 & y < 0 \end{cases} \quad (20)$$

Burada x_i , giriş vektörünün i . elemanını; w_i , i . giriş elemanın ağırlığını; n , giriş sayısını; $F(y)$, transfer (aktivasyon) fonksiyonunu göstermektedir. Girdiler nöronlara gelen verilerdir. Girdiler yapay sinir hücresine başka bir diğer hücreden gelebileceği gibi doğrudan dışarıdan gelen giriş verisi de olabilir. Ağırlıklar girdilerden üretilecek çıktılar üzerindeki etkisini belirler. Toplama fonksiyonu, girdiler ile ağırlıkları çarpan ve bu çarpımları toplayarak hücrenin net girdisini hesaplayan fonksiyondur. Farklı türden toplama fonksiyonları ağırlık mimarisini belirlerken kullanılabilir. Eşitlik 20'de bir toplama işlevi gösterilmiştir.

Transfer fonksiyonu hücreye gelen net girdi üzerinde işlem yaparak hücrenin bu girdiye karşılık üreteceği çıktıyı hesaplar. Bu fonksiyon genellikle doğrusal olmayan bir fonksiyon seçilir. Eşitlik (20)'de bu fonksiyon eşik değeri 0 olan bir adım fonksiyonudur. Hiperbolik tanjant ve sigmoid fonksiyonları diğer yaygın kullanılan transfer fonksiyonlarıdır. Transfer fonksiyonunun çıkışı başka bir hücrenin girişi olarak kullanılabilir gibi doğrudan YSA çıkışı da olabilir.

Sinir hücresinin ağırlıkları ağırlık eğitimi aşamasında, önceden belirlenen bir ölçütü sağlaması için bir eğitim algoritması kullanılarak değiştirilir. Öğrenme işlemi, ağırlık

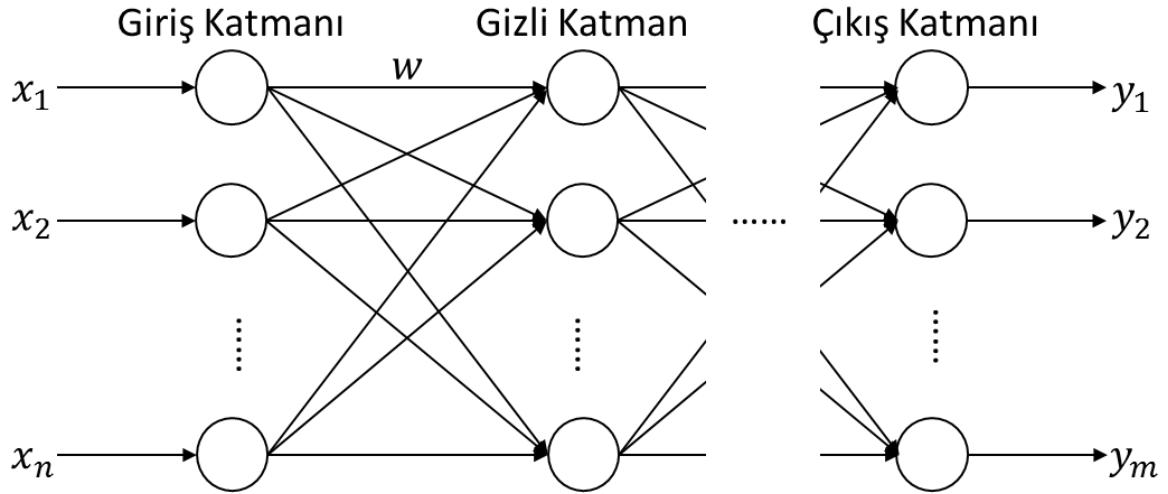
istenilen sonucu üretmesini sağlayan w_i ağırlık vektörünü bulmak için adım adım yaklaşması olarak ifade edilebilir.

Yapay Sinir Ağlarında Öğrenme

YSA'da giriş ve çıkış arasındaki ilişkiyi istenilen sonucu verecek şekilde oluşturmak için değişik eğitim algoritmaları kullanılmaktadır. Bu algoritmalar öğrenme yöntemlerine göre farklı sınıflara ayrılırlar [174]. Konunun başında da belirtildiği gibi danışmanlı öğrenme ve danışmansız öğrenme en yaygın kullanılan iki kategoridir [172]. Genel olarak bir YSA öğrenmesi, danışmandan aldığı sınıf etiketini kullanarak bu sonucu elde edinceye kadar içyapısını uyarlaması olarak tanımlanır.

Danışmanlı öğrenme yönteminde YSA kendi aktif çıkışı ile istenilen çıkış arasındaki hatayı azaltacak şekilde parametrelerini ayarlamaktadır. Bu tür bir öğrenmede girişlerin hangi çıkışı üretmesi gerektiği önceden bilinmektedir. Literatürde birçok öğrenme algoritması ve bunların türevleri vardır [173,175]. Bir YSA'nı eğitmedeki nihai amaç eğitim setini mükemmel bir şekilde öğrenmeye zorlamak değildir. Bunun yerine eğitim süresince rastlamadığı bir giriş değeri için doğru çıkışı üretebilecek iyi bir genelleme yeteneğine sahip olması istenir.

Çok katmanlı ağ (ÇKA – Multi-layer perceptron) danışmanlı öğrenme yöntemini kullanan ağlardan en yaygın kullanılanıdır [173]. ÇKA her biri farklı sayıda nörona sahip artarda dizilmiş katmanlardan oluşur. İlk katmandaki birimler dış dünyadan girişleri alır ve gizli katmandaki birimler ile birebir bağlıdır. Bu katmanda dışarıdan alınan giriş sayısı kadar düğüm olmasına rağmen genelde girdiler herhangi bir işlem yapmadan alt katmanlara iletilirler. Gizli katmandaki birimler bir sonraki katman birimleri ile (çıkış katmanı veya sonraki gizli katman) tam bağlıdır. Gizli katman sayısı ağın yapısına göre değişiklik gösterebilir. Bazı YSA'larda gizli katman bulunmadığı gibi bazılarında ise birden fazla gizli katman bulunabilir. Gizli katmandaki düğüm sayısı giriş ve çıkış katmanlarındaki düğüm sayısından bağımsızdır. Birden fazla gizli katman olan ağlarda gizli katmanlardaki düğüm sayıları da birbirinden farklı olabilir. Çıkış katmanındaki birimler ise ağın çıktısını üretirler. Şekil 25'de ÇKA'nın genel yapısı gösterilmiştir.



Şekil 25. Çok katmanlı ağ

Geri yayılım algoritması (backpropagation) [176] ÇKA'nın eğitiminde yaygın olarak kullanılan algoritmalardan biridir. Bu algoritma hata fonksiyonunu en küçük değerine indirmek için ÇKA'daki bağlantıların ağırlıklarını ayarlamak için kullanılan bir optimizasyon algoritmasıdır [175,177]. Hata fonksiyonu olarak genellikle eşitlik (21)'de verilen Ortalama Kare Hata (Mean Square Error) fonksiyonu kullanılır.

$$E_{MSE} = \frac{1}{n} \sum_{i=1}^n (t_i - y_i)^2 \quad (21)$$

Burada t hedef çıktıyı, y ağın çıktısını ve E_{MSE} hata fonksiyonunu göstermektedir. Çıkış düğümlerinden hata hesaplanır ve sonra diğer katmanlardaki düğümlere geri yayılır. Hatayı en düşük değerine getirmek için geri yayılma sürecinde her bir ağırlık güncellenir.

Geri yayımlı öğrenme genellikle transfer (aktivasyon) fonksiyonu olarak türevi alınabilecek bir fonksiyonu (sigmoid fonksiyonu) kullanır. Öğrenme fonksiyonu olarak eşitlik (22)'de verilen delta fonksiyonu kullanılmaktadır [172].

$$w_{i,j}(yeni) = w_{i,j}(eski) + (\mu [t - f(y_{in})]f'(y_{in})) \quad (22)$$

Burada türevi alınabilen bir fonksiyon kullanılması önemlidir. Çünkü türev bir eğri üzerindeki değişim olarak ifade edilir. Diğer bir deyişle, hatanın en düşük değere indirilmesi demek hatanın türevinin 0 olması anlamına gelir. Geri yayımlı öğrenmede hatanın türevini

ağırlıklar üzerinden öğrenir ve her bir döngüde hata sıfıra doğru yakınsar. Burada μ öğrenme oranını, y_{in} transfer fonksiyonu girişini, f transfer fonksiyonunu ve f' transfer fonksiyonunun türevini göstermektedir. Geri yayılım ağında kullanılan sigmoid (bipolar sigmoid) fonksiyonu ve türevi eşitlik (23) ve (24)'te verilmiştir [172].

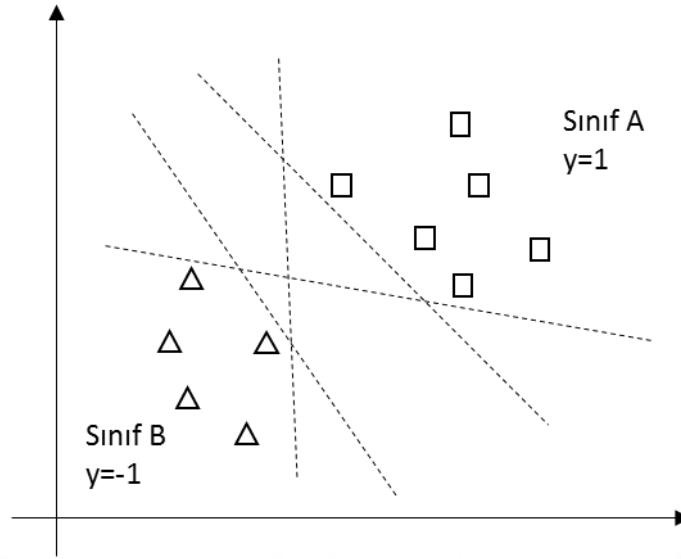
$$f(y) = \frac{2}{1 + e^{-y}} - 1 \quad (23)$$

$$f'(y) = \frac{1}{2}[1 + f(y)][1 - f(y)] \quad (24)$$

1.8.3. Destek Vektör Makinesi

Destek Vektör Makinesi (DVM) veri madenciliğinde sınıflandırma problemlerinde yaygın olarak kullanılan yöntemlerden biridir. DVM, dağılımdan bağımsız olarak çalışabilmektedir ve istatistiksel öğrenme teorisine göre danışmanlı veya yarıdanışmanlı olarak sınıflandırma ve regresyon işlemlerini gerçekleştirebilmektedir. DVM, yüksek boyutlu ve küçük sayıda eğitim veri setinden öğrenebilen istatistiksel öğrenme teorisi çatısı altında yönlendirilmiş yeni nesil bir öğrenme metodudur [178]. Yöntem Vapnik ve Lerner [179] tarafından 1963 yılında geliştirilmiş ve günümüzde kullanılan standart halini ise Cortes ve Vapnik [180] önermiştir.

DVM sınıflandırma işleminde bulunmaya çalışılan ayırıcı düzlem, sınıflandırma probleminde eğitim verisini hatasız bir şekilde ayırabilecek doğrusal bir fonksiyondur [178,181]. Şekil 26'dan görüleceği gibi veriler birbirinden farklı biçimlerde doğrusal olarak ayrılabilirler. Veriyi birbirinden ayıran bu doğrulardan hangisinin seçileceği araştırılması gereken bir konudur. Birbirinden en uzak iki ayırıcı düzlemi elde etmek en uygun çözümü sağlayacaktır.



Şekil 26. İki boyutlu uzayda doğrusal olarak ayrılabilen veriler

İçeriğinde n tane örnek veri bulunan bir eğitim verisi eşitlik (25)'teki gibi gösterilsin.

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), x \in R^d, y \in \{-1, +1\} \quad (25)$$

Bu eğitim verileri uygun w ve b katsayıları ile eşitlik (26)'daki ayırıcı düzlem karar fonksiyonu yardımı ile ayrılabilir.

$$H_0: (W^T X) + b = 0 \quad (26)$$

Bu ifade şu şekilde yazılabilir;

$$\sum_{i=1}^n w_i \cdot x_i + b = 0 \quad (27)$$

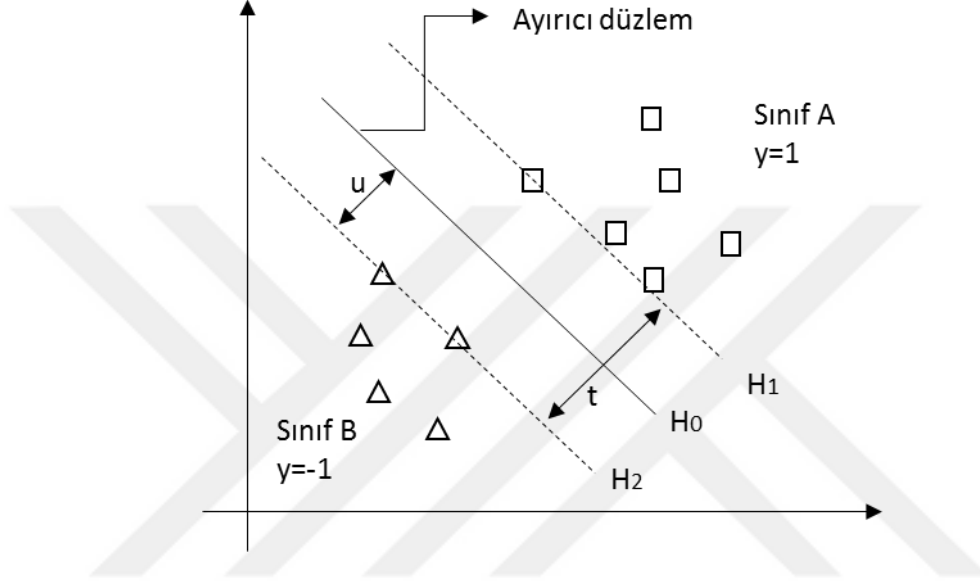
Ayırıcı düzlem n tane veri setinin ayrılmasını sağlayacak şu koşulları yerine getirmektedir. Burada $W = \{w_1, w_2, \dots, w_n\}$ ağırlık vektörünü göstermektedir ve ağırlık değerinin alabileceği sonsuz sayıda farklı çözüm değeri bulunmaktadır. b ise sabit katsayıdır.

$$\begin{aligned} w_i \cdot x_i + b &\geq +1, & \text{eğer } y_i &= +1 \\ w_i \cdot x_i + b &\leq -1, & \text{eğer } y_i &= -1, \quad i = 1, 2, \dots, n \end{aligned} \quad (28)$$

Yukarıdaki eşitlik (28)'i birleştirerek tek bir eşitlik halinde yazabiliriz.

$$y_i[w_i \cdot x_i + b] - 1 \geq 0, \quad i = 1, 2, \dots, n \quad (29)$$

Verilen eğitim veri seti için mümkün olan tüm ayırıcı düzlemler bu eşitlik ile temsil edilebilir.



Şekil 27. Doğrusal olarak birbirinden ayrılabilen eğitim verileri için muhtemel ayırıcı düzlem.

Şekil 27'de H_1 ve H_2 ayırıcı düzlemleri göz önüne alındığında bu düzlemler üzerindeki gözlemler destek vektörler olarak adlandırılır. Bir destek vektör ile çözümü sağlayacak ayırıcı düzlem arasındaki uzaklık eşitlik (30) ile verilir.

$$u = \frac{|(W^T x_1) + b|}{\|w\|} = \frac{1}{\|w\|} \quad (30)$$

Bu formül ile x_1 destek vektörünün H_0 ayırıcı düzleme olan uzaklığı hesaplanır. Bu durumda H_2 ayırıcı düzlemine olan uzaklık ise;

$$t = 2 \cdot u = \frac{2}{\|w\|} \quad (31)$$

olarak hesaplanır. H_1 ve H_2 ayırıcı düzlemleri arasındaki mesafeyi (t) maksimize etmek amaçlandığına göre $\|w\|$ ifadesinin minimize edilmesi t 'yi maksimize edecektir. Daha büyük sınır, sınıflar arasında daha fazla ayırma yeteneğine sahiptir. Sınır üzerinde oluşan veri noktalarına destek vektörler denir. Destek vektörler karar yüzeyine en yakın veri noktaları olduğu için sınıflandırıcıyı en çok zorlayan ve karar yüzeyinin yapısını belirleyen veri noktalarıdır [178].

$$\|w\| = w^t w = w \cdot w = \sum_{i=1}^n w_i^2 \quad (32)$$

eşitlik (32)'deki ifadenin minimizasyonu ayırıcı düzlemi maksimize edecektir.

Doğrusal olarak ayrılabilir durumlar için en iyi ayırıcı düzlemin bulunması doğrusal sabitli ikinci dereceden bir optimizasyon problemidir ve aşağıdaki şekilde ifade edilir.

$$\frac{w^t w}{2}, \quad (33)$$

eşitliğini

$$y_i[w_i \cdot x_i + b] \geq +1, \quad i = 1, 2, \dots, n \quad (34)$$

eşitliğine uygun olarak minimum yapan w ve b değerleri hesaplanır. Burada (33) eşitliğine eklenen $\frac{1}{2}$ çarpanı matematiksel işlemleri kolaylaştırmak için eklenmiştir. Bu optimizasyon problemi ikinci dereceden programlama tekniğinin kullanılmasıyla çözülebilir. Eşitlik (34)'te belirtilen optimizasyon problemi Kurush-Kuhn-Tucker teoremi ölçeklemelerle ikili optimizasyon problemi üretir. Bu yüzden çok yüksek boyutlu ve örnek boyutlarının değişebildiği problemlerde çözüme ulaşmak için standart ikinci dereceden optimizasyon teknikleri kullanılabilir. Bu optimizasyon problemi aşağıdaki şekilde ikili forma dönüştürülür.

İlk olarak Lagrange çarpanları kullanılır.

$$Q(w, b, \alpha) = \frac{1}{2}ww - \sum_{i=1}^n \alpha_i \{y_i(wx_i + b) - 1\} \quad (35)$$

Burada α_i değerleri Lagrange çarpanlarıdır. Bu fonksiyon w ve b açısından düşünüldüğünde minimumlaştırılmalı, $\alpha \geq 0$ açısından düşünüldüğünde ise maksimumlaştırılmalıdır [178]. Kurush-Kuhn-Tucker koşulları uygulanarak w ve b parametreleri α_i cinsinden ifade edilmesi sağlanır. Böylece eşitlik (35)'te verilen formül α_i Lagrange katsayısının maksimumlaştırılması istenen ikili probleme dönüşür. (35) eşitliğinin kısmi türevleri alınarak sıfıra eşitlenir.

$$\frac{\partial Q(w, b, \alpha)}{\partial w} = 0, \quad (36)$$

$$\frac{\partial Q(w, b, \alpha)}{\partial b} = 0, \quad (37)$$

Bu türevlerin çözülmesi ile ayırıcı düzlem ile aşağıdaki özellikleri vermektedir;

$$w = \sum_{i=1}^n y_i \alpha_i x_i = 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, n, \quad (38)$$

$$\sum_{i=1}^n y_i \alpha_i = 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, n \quad (39)$$

Bu koşulların yanı sıra α_i parametreleri bir koşulu daha sağlamalıdır. Eğer örnek veri (x_i, y_i) destek vektör noktalarından birisi ise α_i sıfırdan farklı bir değere sahip olmalıdır. Bunun formül ile ifadesi aşağıdaki şekildedir.

$$\alpha_i \{y_i(wx_i + b) - 1\} = 0, \quad i = 1, 2, \dots, n, \quad (40)$$

Destek vektörler için α_i sıfırdan farklı bir değer alır. İkili problemin yapılandırılması için Lagrange fonksiyonu şu şekilde ifade edilebilir.

$$Q(w, b, \alpha) = \frac{1}{2}ww - \sum_{i=1}^n \alpha_i \{y_i(wx_i + b) - 1\} \quad (41)$$

$$= \frac{1}{2}ww - \sum_{i=1}^n \alpha_i y_i x_i w - b \sum_{i=1}^n \alpha_i y_i + \sum_{i=1}^n \alpha_i \quad (42)$$

Sonuç olarak aşağıdaki eşitlik elde edilir.

$$Q(w, b, \alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i x_j) \quad (43)$$

Eşitlik (43) ikili optimizasyon problemi fonksiyonudur. Bu fonksiyon α_i , $i = 1, 2, \dots, n$ parametrelerine göre maksimumlaştırılmalıdır.

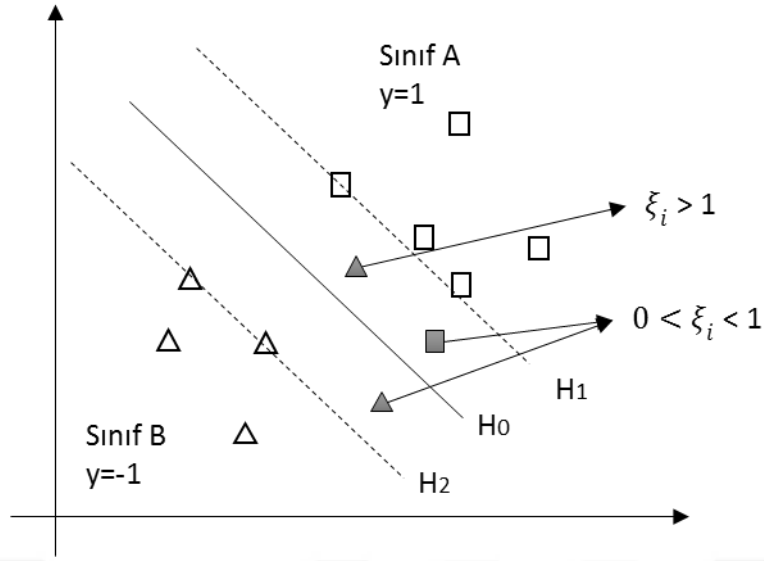
Verilerin Doğrusal Olarak Ayrılamama Durumu

Eğitim verilerinin hatasız olarak ayrılamadığı durumlarda en az hata ile ayrılması tercih edilir ve veri örneklerini doğrusal yapıda ayıran en iyi ayırıcı düzlem kullanılmak istenmektedir. Ayırıcı düzlem yapısı hatadan dolayı (29) eşitliğini sağlamadığı için ayırıcı fonksiyon tarafından ayrılamamaktadır.

Bu durumda ayırıcı düzlemin koşulları tanımlanırken hataları ifade eden ξ_i , $i = 1, 2, \dots, n$ pozitif gevşek değişkeni koşul fonksiyonuna eklenerek probleme çözüm aranır.

$$y_i[w_i \cdot x_i + b] \geq 1 - \xi_i \quad (44)$$

ξ_i gevşek değişkeni bir x_i eğitim örneği için uygun olan sınıfın sınırdan olan izin verilen sapma miktarını göstermektedir. Bu tanıma göre gevşek değişkenlerin değerleri ayrılmazlık durumuna uygun olarak 0 ile 1 arasındadır ve ancak yanlış sınıflandırma durumunda 1'den daha büyük değer alırlar [178]. Bu durum Şekil 28'de gösterilmiştir.



Şekil 28. ξ_i gevşek değişkenlerinin ifadesi

Bu durumda genelleştirilmiş en iyi ayırıcı düzlem için maksimize edilecek fonksiyon ek bir terime (C) sahip olacaktır. Lagrange fonksiyonu şu şekilde ifade edilir.

$$Q(w, \xi) = \frac{1}{2} w w + C \left[\sum_{i=1}^n \xi_i \right]^k \quad (45)$$

Burada $C > 0$ bir sabit değerdir ve kullanıcı tarafından seçilir [181]. Eğer C küçük ise uygun yerde olmayan birçok örneğe izin verilir. $k=1$ için Lagrange fonksiyonu aşağıdaki eşitlikte verilmiştir.

$$Q(w, \xi) = \frac{1}{2} w w + C \sum_{i=1}^n \xi_i \quad (46)$$

Burada kısıtlama;

$$y_i [w_i \cdot x_i + b] \geq 1 - \xi_i, \quad i = 1, 2, \dots, n, \quad (47)$$

Burada belirtilen C değeri dolaylı olarak ayırıcı düzlemin marginini belirlemektedir. Bu optimizasyon problemi yüksek boyutlu uzayda çözülmek isteniyorsa ikili forma dönüştürülmesi gerekmektedir. Daha önce doğrusal ayrılabilirlikte açıkladığı için

burada tekrar edilmeyecektir. Yukarıdaki optimizasyon problemine ait dual Lagrange fonksiyonu aşağıdaki şekilde formüleştirebilir.

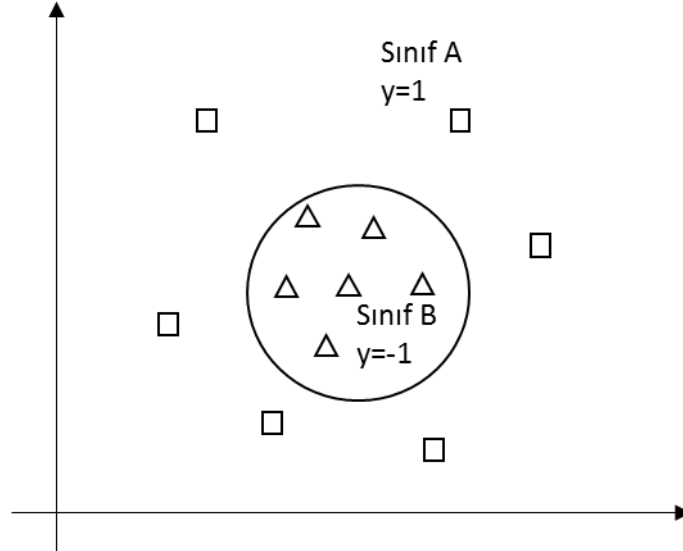
$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i x_j), \quad (48)$$

Burada kısıtlayıcılar;

$$\sum_{i=1}^n y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, n \quad (49)$$

$(x_i, y_i), i = 1, 2, \dots, n$, eğitim verileri ve C sabit bir değerdir.

Şimdiye kadar veri setlerinin bir doğrusal ayırıcı düzlem ile ayrılabilirdiği durumlar incelenmiş ve ayırıcı düzlemlerin iyi bir yaklaşım fonksiyonu olduğu görülmüştür. Çünkü karmaşıklıkları boyuttan bağımsız olarak kontrol edilebilmektedir. Verilerin doğrusal olarak ayrılamadığı durumlarda gevşek değişkenler modele eklenerek çözüm üretildi. Verilerin doğrusal olarak ayrılamadığı durumlarda doğrusal olmayan sınıflandırıcılar kullanılabilir.



Şekil 29. Doğrusal olarak ayrılamayan veri örnekleri

Şekil 29'da görüldüğü gibi iki farklı sınıfa ait verileri doğrusal olarak ayrıştırmak pek mümkün görünmemektedir. Bu durumda farklı bir çözüm bulunması gerekir.

Pratik uygulamaların çoğu doğrusal ayırıcı düzlem kullanılarak gerçekleştirilememektedir. Böyle durumlarda uzaylar arası dönüşüm işlemi yapan bazı dönüşüm fonksiyonları kullanılabilir. Bu örnek verileri doğrusal olarak ayırmak mümkün olmadığı için farklı uzaya dönüşümü doğrusal olmayan yöntemlerin kullanılması gerekmektedir. Bu fonksiyonlardan örneğe uygun olan çekirdek fonksiyonunu iç çarpım kurallarını yerine getirmek için seçilmelidir [178]. Lineer olmayan örnekler için uygulanacak yöntem, $x \in R^n$ gözlem vektörünü daha yüksek mertebeden bir uzayda p vektörlerine dönüştürmek ve bu yeni uzayda doğrusal olarak sınıflandırıcı düzlemleri elde etmektir. Bu p vektörünün yer aldığı özellik uzayı P ile gösterilsin. Bu durumda ϕ ifadesi $R^n \rightarrow R^P$ dönüşümünü yapacak $p = \phi(x)$ olarak ifade edilir. Birden fazla dönüşüm fonksiyonu kullanmak yerine tek bir fonksiyon kullanılabilir. Bu amaçla çekirdek fonksiyonlarından faydalanılır. R^n giriş uzayındaki veri noktaları daha yüksek boyutlu nitelik uzayına haritalandığı için eğitime algoritması sadece P 'deki çarpım noktalarına bağlı olacaktır. Bu çarpım noktaları $\phi(x_i) \cdot \phi(x_j)$ fonksiyonu biçiminde gösterilir.

Geçiş operatörü ϕ 'nin hesaplanması zordur. Bu nedenle $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ yazılabilmesini sağlayan bir ϕ eşleşmesi belirlenebilir. Bu eşleşmeyi sağlayacak K pozitif tanımlı ve simetrik bir çekirdek fonksiyonudur. Aşağıdaki çekirdek fonksiyonları uygulamalarda yaygın olarak kullanılmaktadır.

- Doğrusal çekirdek fonksiyonu: $K(x_i, x_j) = x_i \cdot x_j$
- Polinom çekirdek fonksiyonu: $K(x_i, x_j, c, d) = (c + x_i \cdot x_j)^d$
- Radyal tabanlı çekirdek fonksiyonu: $K(x_i, x_j, \sigma) = e^{-\frac{|x_i - x_j|^2}{2\sigma^2}}$

Optimizasyon için kullanılan Lagrange fonksiyonu aşağıdaki şekildeydi.

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i x_j) \quad (50)$$

Bu fonksiyonda $x_i x_j$ çarpımı yerine daha büyük boyutlu özellik uzayına dönüştüren ϕ fonksiyonu kullanılır. Yeni özellik uzayı için optimizasyon fonksiyonu şu şekilde olacaktır.

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\phi(x_i) \phi(x_j)) \quad (51)$$

Burada $\phi(x_i) \phi(x_j)$ çarpımı yerine eşitlik (52)'deki çekirdek fonksiyonu kullanılabilir.

$$K(x_i, x_j) = \phi(x_i) \phi(x_j) \quad (52)$$

Bu durumda optimizasyon problemimiz son halini almış olacaktır.

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (K(x_i, x_j)) \quad (53)$$

Kısıtlayıcılar;

$$\sum_{i=1}^n y_i \alpha_i = 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, n \quad (54)$$

Burada $H_{ij} = y_i y_j K(x_i, x_j)$ olarak alınırsa,

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha^T H \alpha \quad (55)$$

yazılabilir. Bu ifadeye göre karar fonksiyonu,

$$D(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \quad (56)$$

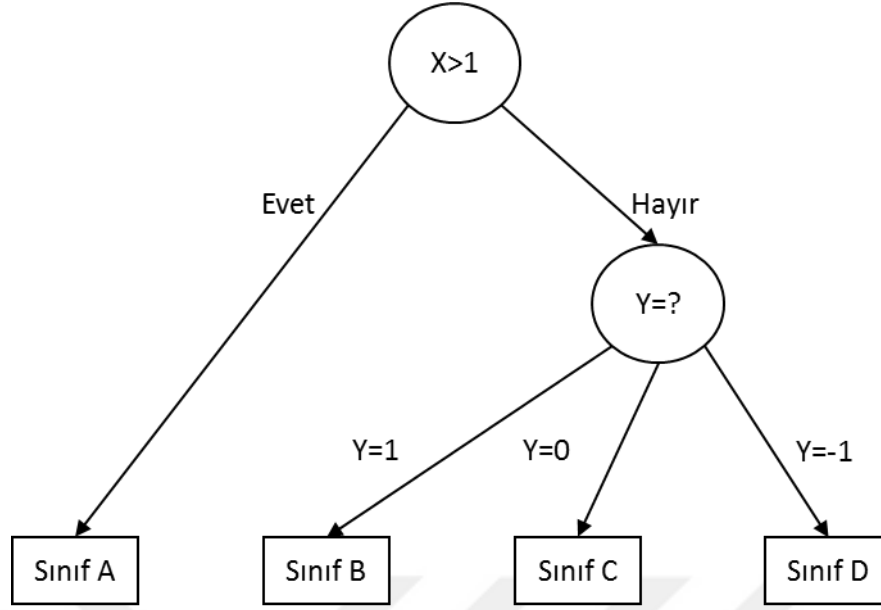
Bu daha önceki karar fonksiyonu ile benzerdir. Tek fark iç çarpımda çekirdek fonksiyonunun gelmesidir. Burada doğrusal ayırıcı düzlem olmadığından b terimi ihmal edilebilir [181].

DVM ile Çok Sınıflı Sınıflandırma

DVM sınıflandırıcılar doğal olarak iki sınıf bulunan durumlar için eğitilmektedirler. Çünkü DVM iki sınıfı birbirinden ayırabilen bir ayırıcı bir düzlem araştırılması üzerine kurulmuş bir yöntemdir. DVM eğitimi için pozitif veya negatif sınıfa ait eğitim verisine eğitim aşamasına ihtiyaç duyulur. Çok sınıflı sınıflandırma durumunu içeren uygulamalarda çözüm için birden çok sayıda ayırıcı fonksiyonun bulunması gerekmektedir. Bu tür sınıflandırma işlemleri için iki yaklaşım benimsenmiştir. Bunlardan ilki bütün sınıfların doğrudan sınıflandırılmasıdır. İkincisi ise ikili sınıflandırmayı kullanarak dolaylı sınıflandırma yapmaktır. Örneğin 1-v-r prensibinde sınıflandırıcı bir örneği +1 olarak sınıflandırdığında geri kalan diğer tüm sınıfları -1 olarak sınıflandırır [178].

1.8.4. Karar Ağaçları ile Sınıflandırma

Karar Ağaçları (KA) sınıflandırma ve tahmin için yaygın olarak kullanılan parametrik olmayan bir veri madenciliği yöntemidir. KA kolay yorumlanabilir ve anlaşılabilir olması açısından diğer sınıflandırıcılara göre daha avantajlı sayılabilir. KA yapısını giriş ve çıkış setlerinden oluşturan danışmanlı bir öğrenme algoritmasıdır. Tipik bir karar ağacı öğrenme sistemi, arama uzayında bir çözümü bulmak için yukarıdan aşağıya arama stratejisini kullanmaktadır. KA öznelikleri test edilmiş düğümlerden oluşur. Tek değişkenli ağaçta her bir iç düğüm için test düğümü test için niteliklerden sadece birini kullanır. Bir düğümün giden dalları düğümden testin tüm olası sonuçlarına karşılık gelmektedir.



Şekil 30. X ve Y değişkenleri için basit bir karar ağacı

Şekil 30'da X ve Y giriş öznitelikleri için örnekleri sınıflandıran basit bir karar ağacı gösterilmiştir. Burada X'in birden büyük olmadığı ve Y'nin bire eşit olduğu tüm örnekler Sınıf B olarak sınıflandırılacaktır. Ancak X'in birden büyük olduğu durumlarda Y'ye bakılmaksızın sınıflandırma sonucu Sınıf A olarak bulunur [163]. KA kök olarak adlandırılan ve herhangi bir girişi olmayan bir düğümlle başlayan yönlü bir ağaçtır. Tüm diğer düğümlerin tek bir girişi mevcuttur. Çıkışı olan düğümler test düğümü olarak adlandırılır. Bunlar haricinde diğer tüm düğümler yaprak veya karar düğümü olarak adlandırılır [182]. Her bir yaprağa en doğru hedef değerini gösteren bir sınıf atanır. Yaprak kesin bir değere sahip hedef özelliği olasılığını belirten bir olasılık vektörü tutar. Bu nedenle, son yapraktan köke kadar en olası yolu bulmak için yaprakların her biri için olasılık değerleri çarpılarak hesaplama yapılır. Ağacın belli dalları kesilerek veya karakteristiği belirleyen özellikleri değiştirilerek hesaplama verimliliği geliştirilebilir. ID3, C4.5, CART, CHAID ve MARS gibi çok yaygın olarak kullanılan KA algoritmaları vardır.

ID3 algoritması ağacın kök düğümünde tüm eğitim örnekleri ile başlar. Bir öznitelik bu örnekleri bölümlenmek için seçilir. Bu özneliğin her bir değeri için bir dal oluşturulur ve dal tarafından belirtilen öznitelik değerine sahip örneklerin ilgili alt kümesi yeni oluşturulan alt düğüme taşınır. Bir düğümdeki tüm örnekler bir sınıfa ait olana kadar algoritma özyinelemeli olarak her alt düğüme uygulanır. Karar ağacında yaprağa kadar giden her bir yol bir sınıflandırma kuralını temsil eder. ID3 ve C4.5 algoritmalarında öznitelik seçimi bir

düğümde bulunan örneklerden hesaplanan bilgi entropisini en aza indirme temeline dayanmaktadır. Bilgi entropisine dayalı yaklaşım bir veritabanındaki bir örneği sınıflandırmayı sağlayacak test sayısını en aza indirmeye çalışır. ID3 öznitelik seçimi kısmı karar ağacının karmaşıklığı verilen özniteliğin değeri tarafından iletilen bilgi miktarı ile güçlü bir ilişki olduğu varsayımına dayanmaktadır. Bilgi tabanlı sezgisel bir yapı en yüksek bilgi kazancını sağlayan özniteliği seçer. Bu öznitelik örneği sınıflandıracak alt ağaç için ihtiyaç duyulacak bilgiyi minimize eder. ID3 algoritmasının genişletilmiş versiyonu olan C4.5 algoritması sınıflandırma domenini kategorik özelliklerden nümerik olanlara genişletmektedir. Bu ölçüt veriyi düşük sınıf entropisine sahip örneklerin çoğunluğunun tek sınıfa ait olduğu alt kümelerle bölen öznitelikleri tercih eder. Algoritma aslında sınıflar arasındaki yerel ayrımı maksimum derecede sağlayan özniteliği seçer [163].

1.9. Sınıflandırma Başarımının Değerlendirilmesi İçin Performans Ölçütleri

Sınıflandırıcıların örnekler için hangi doğrulukta tahmin yaptıklarının ölçülebilmesi için farklı ölçütler kullanılmaktadır. Aşağıdaki tabloda sınıflandırıcı değerlendirme ölçütleri özetlenmiştir.

Çalışmalarda yaygın olarak doğruluk, duyarlılık, özgünlük, kesinlik ve F ölçütleri yaygın olarak kullanılmaktadır. Ancak doğruluk terimi sınıflandırıcının tahmin yeteneğini gösteren genel bir terim olarak kullanılmaktadır. İki sınıflı bir tanıma sistemi düşünüldüğünde sistem 0 veya 1 olarak yani pozitif veya negatif olarak karar verir. Verilen her karar doğru (True) veya yanlış (False) olabilir. Böylece her bir karar için iki tür cevap vardır ve bu karar için dört farklı durumu ortaya çıkarır. Eğer sınıflandırılan örnek pozitif ve pozitif olarak sınıflandıysa doğru pozitif (TP) olarak, sınıflandırma sonucu negatif ise yanlış negatif (FN) olarak adlandırılır. Eğer örnek negatif ve sınıflandırma sonucu negatif ise doğru negatif (TN) olarak, pozitif olarak sınıflandırıldı ise yanlış pozitif (FP) olarak adlandırılır.

Tablo 3. Performans Ölçütleri

Ölçüt	Formül
Doğruluk, Tanıma oranı (accuracy)	$\frac{TP + TN}{P + N}$
Hata Oranı, Yanlış sınıflandırma oranı (error rate)	$\frac{FP + FN}{P + N}$
Duyarlılık, doğru pozitif (TP) oranı (sensitivity, recall)	$\frac{TP}{P}$
Özgünlük, doğru negatif (TN) oranı (specificity)	$\frac{TN}{N}$
Kesinlik (precision)	$\frac{TP}{TP + FP}$
F, F1, F-Skoru	$\frac{2 \times \text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}}$

Tablo 4. Karşıtlık Matrisi

		Sınıflandırma sistemini çıktısı - Tahmin		Toplam
		Doğru	Yanlış	
Gerçek Durum	Doğru	Doğru Pozitif (TP)	Yanlış Negatif (FN)	P
	Yanlış	Yanlış Pozitif (FP)	Doğru Negatif (TN)	N
Toplam		P'	N'	P + N

Karşıtlık matrisi sınıflandırıcının farklı sınıfları nasıl tanıdığıın analizi için faydalı bir araçtır. Bu matriste TP ve TN sınıflandırıcının tanımayı doğru yaptığını, FP ve FN ise yanlış yaptığını gösterir. İki sınıflı sınıflandırıcı için karşıtlık matrisi (hata matrisi) Tablo 4'te gösterilmiştir.

Sınıflandırılacak sınıf sayısı m adet ($m \geq 2$) olduğunda karşıtlık matrisi $m \times m$ boyutlu bir tablodur. Her bir C_{ij} , i . sınıfa ait olup sınıflandırıcı tarafından j . sınıf olarak etiketlenen eleman sayısını gösterir. Sınıflandırıcının sınıflandırma doğruluğu iyi ise ideal olarak örneklerin büyük bir çoğunluğunun bu matrisin köşegenlerinde ($C_{1,1}, \dots, C_{m,m}$) yerleşmiş olması beklenir [183].

Sınıflandırıcının doğruluğu sınıflandırıcı tarafından doğru olarak sınıflandırılan test kümesinin yüzdesidir.

$$\text{Doğruluk} = \frac{TP + TN}{P + N} \quad (57)$$

Örüntü tanıma literatüründe bu değer sınıflandırıcının genel tanıma oranı olarak bilinir ve sınıflandırıcının çeşitli sınıfları ne kadar iyi sınıflandırdığını gösterir.

Bu oranın tam tersine ise hata oranı veya yanlış sınıflandırma oranı denir. Bu değer 1-*doğruluk* eşitliği ile veya eşitlik (58)'deki formül ile hesaplanır.

$$\text{Hata Oranı} = \frac{FP + FN}{P + N} \quad (58)$$

Ana sınıfın daha seyrek olduğu ve sınıfların dengesiz olarak dağıldığı durumu göz önüne alalım. Bu durumda veri seti dağılımı negatif sınıfın önemli bir çoğunluğunu ve pozitif sınıfın daha azını yansıtır. Örnek vermek gerekirse medikal bir veri setinde kanser gibi veri seti içinde daha seyrek olacak ana sınıflar olabilir. Medikal verileri sınıflandırmak için bir sınıflandırıcıyı "kanser" özneteliği için evet veya hayır olarak sınıflandıracak şekilde eğitilmiş olsun. Sistemin doğruluğu %97 olduğunda sınıflandırıcı gayet başarılı bulunabilir. Ana sınıfımız kanser olması durumuydu. %97 kanser olmadığıın sınıflandırıldığı durumda %3'lük diğer kısmın kanser olduğu anlamına gelmeyebilir. Bu durumda %97'lik bir

doğruluk kabul edilemez. Bu durumda sınıflandırıcının pozitif örnekleri (kanser=evet) ve negatif örnekleri (kanser=hayır) nasıl tanıdığını gösteren diğer ölçütlerin hesaplanmasına ihtiyaç duyulur. *Duyarlılık* ve *Özgünlük* bu amaç için kullanılır. Duyarlılık doğru pozitif oranı olarak da bilinir [183]. Bu ölçütler şu şekilde hesaplanır;

$$Duyarlılık = \frac{TP}{P} \quad (59)$$

$$Özgünlük = \frac{TN}{N} \quad (60)$$

Kesinlik ölçütü sınıflandırmada yaygın olarak kullanılır. Pozitif olan örneklerin hangi yüzde ile pozitif olarak etiklendiğini belirtir. Kesinlik ölçütü eşitlik (61)'deki formül ile hesaplanır.

$$Kesinlik = \frac{TP}{TP + FP} \quad (61)$$

Kesinlik ve Duyarlılığı kullanmanın farklı bir yolu bu iki değeri tek bir ölçütte birleştirmektir. Bu bazı kaynaklarda F1 veya F-Skoru olarak bilinen F ölçütü yaklaşımıdır ve eşitlik (62)'deki formül ile hesaplanır.

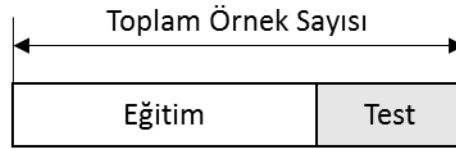
$$F = \frac{2 \times Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (62)$$

F ölçütü kesinlik ve duyarlılığın harmonik ortalamasıdır. Kesinlik ve duyarlılığa eşit ağırlık verir.

Özetlemek gerekirse doğruluk ölçütü veri sınıfları düzgün bir şekilde dağılmış ise doğru olarak çalışır. Diğer ölçütler sınıfların sayısının düzgün dağılmadığı duruma, ana sınıfın daha nadir rastlandığı durumlara daha uygundur.

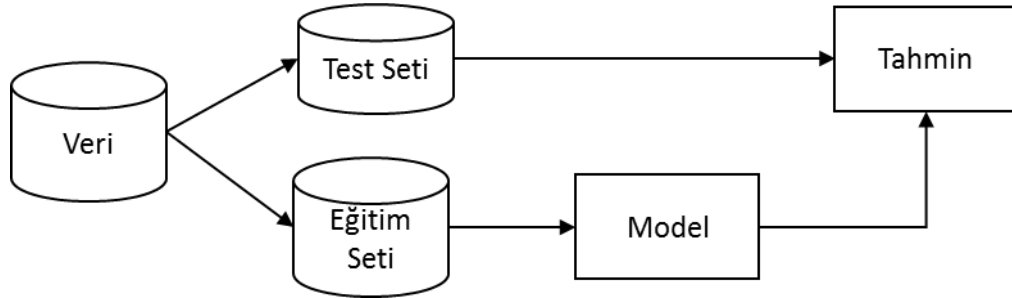
1.10. Sınıflandırma İçin Veri Setinin Organizasyonu

Güvenilir sınıflandırma sonuçlarının elde edilebilmesi için sınıflandırma ve test aşamalarında kullanılacak veri setlerinin nasıl organize edildiği çok önemlidir. Yaygın kullanılan sınıflandırma yöntemi veri setinin eğitim ve test olarak birbirinden bağımsız ve rasgele iki ayrı alt kümeye ayrıldığı yöntemdir (Şekil 31).



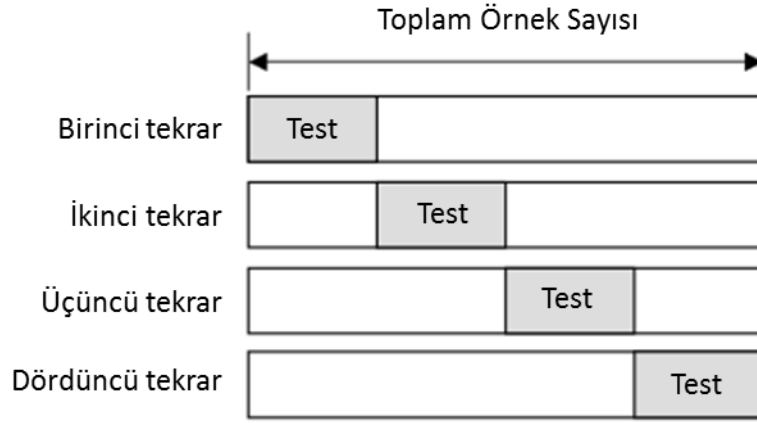
Şekil 31. Veri kümesinin organizasyonu

Bu yöntemde genellikle verinin 2/3'ü eğitim alt kümesine ve geri kalan 1/3'ü test alt kümesine atanır. Küme oranlarının %50- %50 olduğu durumlar da sınıflandırmada yaygın olarak kullanılır. Şekil 32'de görüldüğü gibi eğitim seti sınıflandırma modelini eğiterek değerlendirme aşaması için hazır hale getirir ve sonrasında modelin doğruluğu test seti ile tahmin edilir.



Şekil 32. Doğruluğun tahmin edilmesi

Bu yöntemin biraz daha geliştirilmiş hali k-kere çapraz doğrulamadır. k-kere çapraz doğrulamada veri rasgele olarak birbiri ile örtüşmeyen k tane eşit boyutlu alt küme (D_1, D_2, \dots, D_k) ayrılır. Eğitim ve test aşaması k kez tekrar edilir. i . tekrarda D_i alt kümesi test için ayrılır ve geri kalan alt kümeler birleştirilerek eğitimde kullanılır. Örneğin, ilk tekrarda D_2, \dots, D_k birleştirilerek modelin eğitiminde kullanılırken D_1 eğitilen modelin testinde kullanılır. Şekil 33'de $k=4$ için eğitim ve test kümelerinin organizasyonu gösterilmiştir.



Şekil 33. $k=4$ için veri seti

Bir önceki modelden farklı olarak k -kere çapraz doğrulama yönteminde her bir örnek eşit sayıda eğitim ve test verisi olarak kullanılmaktadır. Sınıflandırma işleminde ölçütler hesaplanırken k tekrarda elde edilmiş sonuçların ortalaması alınır. K -kere çapraz doğrulama yönteminin her seferinde bir örneğin test için dışarıda bırakılıp (leave-one-out) diğer örneklerin eğitim için kullanıldığı (k =örnek sayısı) özel bir türü vardır. K -kere çapraz doğrulamada $k=10$ yaygın olarak kullanılır ve 10-kere çapraz doğrulama adını alır.

1.11. Literatür Çalışması

1.11.1. Erken Karıncık Kasılması Virusunun Sınıflandırılması ile İlgili Literatür Değerlendirmesi

Son yıllarda EKK vurularının otomatik olarak algılanmasını sağlayan birçok çalışma yapılmıştır. Araştırmacılar EKK'nı sınıflandırmak için çoğunlukla zaman-frekans analizi teknikleri, istatistiksel ölçüler ve karma yöntemler kullanmışlardır. Oliveira vd. MIT-BIH aritmi veritabanındaki kayıtlar üzerinde dinamik ve statik Bayes ağları kullanılarak EKK sınıflandırmışlar ve dinamik Bayes ağlarının daha iyi başarımlar elde ettiği sonucuna varmışlardır [184]. Ittatur vd. üç zaman domeni özneliğini kullanarak gerçek zamanlı uygulamalarda EKK sınıflandırmaya çalışmışlardır [185]. Bunlar RR-aralığı, QRS-genişliği ve QRS şeklidir. Sınıflandırma aşamasında gömülü sistemler için uygun basit karar kuralları kullanmışlardır. Algoritmalarını MIT-BIH aritmi veritabanından aldıkları 26 ECG işareti üzerinde denemişlerdir. EKK sınıflandıran bir diğer çalışma Bortolan vd. tarafından ortaya

konmuştur. Çalışmada YSA, K-NN, diskriminant analizi (DA) ve bulanık mantık (BM) sınıflandırma yöntemleri MIT-BIH aritmi veritabanından alınan kayıtlar üzerinde kullanılmış ve normal vuru ve EKK vurularının sınıflandırılmasında sınıflandırma yetenekleri karşılaştırılmıştır [186]. Yazarlar çalışmalarında genel ve yerel diye iki veri kümesi oluşturmuş ve testlerini bu veri kümeleri üzerinde yapmıştır. Genel kümede 130 normal, 130 EKK vurusu olmak üzere 260 vuru bulunmaktadır. İki grup yerel veri seti oluşturulmuş. Birinci sette 38 normal ve 38 PVC olmak üzere 76 vuru bulunmaktadır. İkinci grup yerel veri setinde ise test edilen hastaların ilk %30'luk kalp vuruları bulunmaktadır. Genel veri setinde DA ve BM diğer iki sınıflandırıcıya göre daha iyi sonuçlar elde etmiştir. Yerel setteki ilk grupta K-NN algoritması ile en iyi sonucu elde etmişken yerel setteki ikinci grupta YSA en iyi sonucu elde etmiştir. Ebrahimzadeh ve Khazae YSA kullanarak EKK'nı algılamaya çalışmışlardır. EKG'deki gürültünün azaltılması için dalgacık dönüşümü kullanılmıştır. Öznitelik çıkarma aşamasında 10 yapısal özneliğin yanında bir tane zaman aralığı özneliği kullanılmıştır. MIT-BIH aritmi veritabanından alınan 7 işaret dosyası kullanılarak çok katmanlı YSA ile sınıflandırma yapılmıştır. Li vd. EKK algılama için düşük karmaşık ve veri uyarlamalı bir yöntem tasarlamış ve olasılık temelli bir yöntem geliştirmiştir [187]. EKK algılayıcı, iki şablon eşleştirme prosedüründen elde edilen korelasyon katsayıları analiz edilerek geliştirilmiştir. Yaklaşımlarını test etmek için MIT-BIH aritmi veritabanından alınan 22 işaret kullanılmıştır. Ribeiro vd. EKK algılama için manifold öğrenme üzerine yoğunlaşmış ve sınıflandırma için DVM kullanmıştır [188]. Bir diğer çalışmada Foo vd. YSA tabanlı bir EKG örüntü tanıma yöntemi önermiş ve MIT-BIH aritmi veritabanından alınan kayıtlar üzerinde iki katmanlı geri beslemeli bir YSA'yı vuruları sınıflandırmak için kullanmıştır [189]. Bir diğer çalışmada [190] yazarlar EKK vurusunu sınıflandırmıştır. EKG'deki şekilsel özellikleri elde etmek için dalgacık dönüşümünü ve sınıflandırma için YSA kullanılmıştır. Bir diğer çalışmada Jenny vd. EKK sınıflandırmak için k-ortalama ve bulanık C-ortalama algoritmalarını kullanmıştır. Öznitelik çıkarımı için BBA algoritmasından faydalanmışlardır [191]. Talbi vd. yaptıkları çalışmada EKK algılama için kesirli lineer tahmin yöntemini kullanmışlardır. Bu yöntemde QRS kompleksleri 3 katsayı olarak ifade edilmiştir. Çalışmada hesaplanan bu katsayılar ile doğrusal tahmin katsayılarının sınıflandırma başarımları karşılaştırılmıştır [192].

1.11.2. EKG Vurularının Sınıflandırılması ile İlgili Literatür Değerlendirmesi

Araştırmacıların son yıllarda yaptıkları çalışmalarda matematiksel modeller ve istatistiksel analizleri kullanan uygulamalar ile fizyolojik kardiyak olayların yorumlanması için birçok avantajlı çözümler sunmuştur. Liu vd. EKG vurularının sınıflandırılması için kendi geliştirdikleri, k-ortanca ve k-ortalama algoritmalarını içeren sözlük öğrenme yöntemini kullanmıştır. Çalışmada MIT-BIH’te bulunan farklı veri tabanlarından alınan sekiz sınıfa ait 1200 vuru test için kullanılmış ve öznitelik çıkarımı için vektör kuantalama algoritmasından faydalanılmıştır [193]. Thomas vd. EKG aritmilerinin sınıflandırılması için dalgacık dönüşümleri ile elde edilmiş öznitelikleri kullanmıştır. Çalışmada sınıflandırıcı olarak çok katmanlı YSA kullanılmıştır. Önerilen öznitelik kümesinin performansı QRS kompleksinden çıkarılan ayırık dalgacık dönüşümü ile elde edilmiş alt bant öznitelikleri ile karşılaştırılmıştır. Çalışmada MIT-BIH aritmi veri tabanından alınan veri kümesi kullanılmış ve biri normal, dördü anormal olmak üzere beş ayrı sınıfta vuru sınıflandırma işlemi gerçekleştirilmiştir [194]. Dilmak ve Korurek yaptıkları çalışmada değiştirilmiş yapay arı kolonisi algoritmasını EKG vurularının sınıflandırılmasında kullanmışlardır. Çalışmada öznitelik olarak QRS kompleksine ait morfolojik özelliklerin yanı sıra kalp vuru süreleri de kullanılmıştır. MIT-BIH aritmi veritabanından alınan üç vuru türünde sınıflandırma gerçekleştirilmiş ve elde edilmiş sonuçlar literatürde kullanılan birçok sınıflandırıcı sonucu ile karşılaştırılmıştır [195]. Bir diğer çalışmada Nazarahari vd. EKG vurularının sınıflandırılmasında öznitelik çıkarımı için yeni dalgacık fonksiyonlarını kullanmışlardır. Sınıflandırma algoritması olarak çok katmanlı YSA kullanılmıştır. Çalışmada özniteliklerin azaltılması için TBA kullanılmış, biri normal toplam altı vuru türü sınıflandırılmış ve sonuçları karşılaştırılmıştır [196]. Çalışma [197]’de yazarlar EKG aritmilerinin sınıflandırılması için model önermiştir. Çalışmada özniteliklerin hesaplanması için spektral korelasyon ve TBA analizi kullanılmış, sınıflandırıcı olarak DVM kullanılmıştır. MIT-BIH aritmi veritabanından alınan beş ayrı vuru türü deneylerde ayrıştırılmıştır. Escalona-Morán vd. YSA’ya benzer bir hesaplama altyapısı olan Rezervuar Hesaplamayı kullanmışlardır. Yazarlar geliştirdikleri yöntemlerin EKG işaretlerinin ön işleme aşamalarında daha hızlı çalıştığını iddia etmektedir. MIT-BIH aritmi veritabanından aldıkları kayıtlarda Tıbbi Alet Geliştirme Derneği (AAMI) tarafından sunulan rehber kullanılarak sınıflandırma işlemi gerçekleştirmiştir [198]. Oster vd. karıncık kaynaklı aritmilerin sınıflandırılması için yarı-danışmanlı bir yöntem önermiştir. MIT-BIH veritabanındaki karıncıkla ilgili vuru türlerinin

algılanması için değiştirilmiş Kalman Filtresi kullanılmıştır [199]. Bir diğer çalışmada yazarlar EKG aritmilerinin sınıflandırılması için Sürekli Dalgacık Dönüşümü temelli bir yöntem önermiştir ve GA kullanılarak yeni bir DD temelli sınıflandırıcı geliştirilmiştir. Sınıflandırma için AAMI standardında etiketler kullanılmıştır [200]. Alshraideh vd. EKG işaretlerinden elde edilen özniteliklere ek olarak hastanın demografik özelliklerini de kullanan bir sistem önermiştir. Web tabanlı ve uzaktan algılama yapmayı sağlayan bir sistem üzerinde çalışan yazarlar karşılaştırma için farklı sınıflandırma algoritmalarının sonuçlarını irdelemiştir [201]. Bir diğer çalışmada yazarlar Stockwell dönüşümü kullanarak EKG aritmilerinin sınıflandırılması için bir yöntem önermiştir. Çalışmada S-Dönüşümü şekilsel özniteliklerin çıkarılması için kullanılmış ve bu öznitelikler geçici özniteliklere eklenmiştir. Sınıflandırıcı olarak YSA ve DVM kullanılmıştır [202]. Sumathi vd. ANFIS kullanarak EKG vurularının sınıflandırılması için karma bir yöntem önermiştir. Bu yöntem Symlet DD temelinde EKG işaretlerini analiz eden ve tehlikeli kalp hastalıkları ile ilişkili parametreleri çıkaran bir yöntemdir. Bu parametreler ANFIS sınıflandırıcıda kullanılarak önemli beş EKG işaret türü sınıflandırılmıştır [203].

Yukarıda belirtilen çalışmalar alanda önemli dergilerde son birkaç yılda (2014-2017) yayınlanmış çalışmalardır. Bunlara ek olarak bu alanda en çok başvurulan çalışmalar aşağıda incelenmiştir. Ceylan vd. EKG aritmilerinin sınıflandırılması için otomatik bir tanı yöntemi geliştirmiştir. Önerilen metotta tip-2 bulanık c-ortalama ve YSA birleştirilmiş ve yeni bir yöntem elde edilmiştir. Çalışmada EKG vurularının sınıflandırılması için sınıflandırma doğruluğunun artırılması amaçlanmıştır ve MIT-BIH aritmi veritabanından alınan kayıtlar üzerinde deneyler yapılmıştır [204]. Melgani vd. yaptıkları çalışmada EKG vurularının otomatik tanısı için DVM kullanmışlardır. Buna ek olarak Parçacık Sürüsü Optimizasyonu yöntemi kullanılmış ve DVM algoritmasının genelleştirme yeteneği geliştirilmiştir [205]. Bir diğer çalışmada yazarlar kişisel EKG vuru tanıma işlemi için blok-tabanlı YSA kullanmıştır. Öznitelik olarak yazarlar Hermite dönüşümü katsayılarına ek olarak R-R aralığını kullanmıştır [206].

Güler ve Übeyli yaptıkları çalışmada EKG vurularının sınıflandırılması için birleşik YSA mimarisini kullanmıştır. Çalışmada EKG işaretleri DD kullanılarak zaman-frekans gösterimi olarak ayrıştırılmış ve dağılımı temsil edecek istatistiksel değerler hesaplanmıştır. Çalışmada Physiobank veritabanından alınan dört vuru türü sınıflandırılmıştır [207]. Bir diğer çalışmada yazarlar MIT-BIH aritmi veritabanından alınan 22 kayıta bulunan toplam

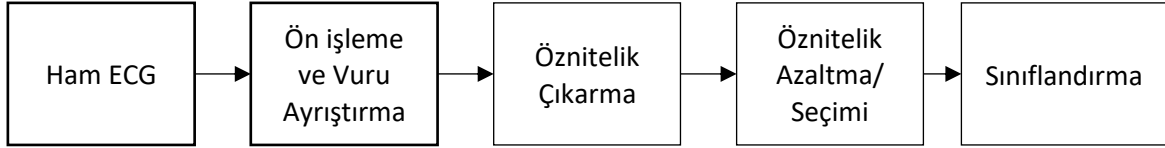
50000 vurudan oluşan veri kümesi üzerinde vuru sınıflandırma işlemi gerçekleştirmiştir. Sınıflandırmada öznitelik olarak EKG şekil bilgisi, kalp vuru aralıkları ve RR aralıkları kullanılmıştır. Sınıflandırmada sınıf etiket bilgisi olarak AAMI standartlarındaki sınıf etiketleri kullanılmıştır [36]. Engin yaptığı çalışmada EKG vurularının sınıflandırılması için yapay-bulanık ağları kullanmıştır. EKG vurularının sınıflandırılmasında EKG vurusu yerine vuruyu temsil edecek, otoregresif model katsayıları, yüksek-mertebeden kümülant ve DD varyanslarını öznitelik olarak kullanılmıştır. MIT-BIH aritmi veritabanından alınan dört ayrı sınıfa ait veriler sınıflandırma deneylerinde kullanılmıştır [208]. Benzer şekilde yapılan bir diğer çalışmada yazarlar EKG vurularının sınıflandırılması için bütünleşik YSA kullanmıştır. Çalışmada bir kalp vurusuna ait Fourier analizi ve DD hesaplamaları öznitelik olarak kullanılmıştır. EKG öznitelikleri ıraksaklık değerine göre dinamik programlama yöntemleri ile tespit edilmiştir [4]. Osowski vd. yaptığı çalışmada EKG vurularının sınıflandırılması için bulanık bütünleşik YSA kullanmıştır. Çalışmada öznitelik seçimi için ikinci, üçüncü ve dördüncü dereceden kümülantlar kullanılmış ve c-kümeleme ve Gustafson-Kessel algoritmaları YSA'nın kendi kendini organize etmesi için kullanılmıştır [209].

2. YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME

Bu çalışmada ilk olarak, erken karıncık kasılması (EKK) virusunun literatürdeki diğer EKK algılayıcılardan daha yüksek bir doğruluk ile algılayacak bir sınıflandırıcı modeli geliştirilmiştir. EKK vurularının sınıflandırılması için yeni öznitelikler üretmek için işaretin bir kalp vurusuna karşılık gelen zaman serisi ham veri olarak ele alınmıştır. Kullanılan bu veriler gürültü dahil birçok bilgi içerdiği için boyut indirgeme yöntemleri ile ayrıştırmaya çalışılmış ve bu amaçla TBA, BBA ve ÖH yöntemleri kullanılmıştır. Zaman serilerinden bir vuruyu temsil edecek en iyi alt kümenin belirlenmesi için GA'lerden faydalanılmıştır. Çalışmada zaman serilerinin ve GA ile azaltılmış alt kümelerinin öznitelik olarak EKK vurularının sınıflandırılmasında kullanılması, yapılmış diğer çalışmalardan farklılık göstermekte ve daha iyi sonuç vermektedir [210].

Daha fazla vuru tipinin sınıflandırılabilmesi için çalışma genişletilmiştir. Bu aşamada kullanılan veritabanında bulunan en yaygın dokuz vuru türü belirlenmiş ve deneylerde bu dokuz vuru türü sınıflandırmak için testler yapılmıştır. Bu vuruları en başarılı şekilde sınıflandırmak için yeni öznitelikler hesaplanmıştır. Sınıflandırma algoritmalarının başarımları test edilmiş ve bu başarımları elde ettikleri sınıflandırma parametreleri irdelenmiştir.

Kalp vuru sınıflandırması yapan sistemin genel şeması Şekil 34'te gösterilmiştir. İlk aşamada ön işleme adımı vardır. Bu adımda taban hattı sapma ve güç hattının etkisinden kaynaklanan gürültüler giderilmektedir. QRS algılama ve vuru ayrıştırma aşamaları da bu adımda gerçekleştirilmektedir. Öznitelik çıkarma aşamasında ayrıştırılmış her bir vurudan vuruyu temsil edecek öznitelikler hesaplanmaktadır. Bu öznitelikler işarete ait istatistiksel ve matematiksel özniteliklerdir. Öznitelik azaltma aşamasında TBA, BBA ve ÖH kullanılmıştır. Öznitelik seçimi için GA ve sıralı arama algoritmaları kullanılmıştır. Sınıflandırma algoritması olarak K-NN, YSA, DVM ve KA kullanılmış ve her birinin sınıflandırma başarımları karşılaştırılmıştır [27].



Şekil 34. Vuru sınıflandırma sisteminin genel blok diyagramı

Çalışmada kullanılan veriler, EKG konusunda çalışan araştırmacılar için standart haline gelen MIT-BIH Aritmi Veritabanından alınmıştır [14,37]. Geliştirilen sistemde bulunan sınıflandırıcıların birbirleri ve literatürdeki diğer çalışmaların performansları ile karşılaştırılması için doğruluk, duyarlılık, özgünlük ve kesinlik performans kriterleri kullanılmıştır. Kullanılan veritabanı bilgileri ve Şekil 34'te verilen blok diyagramdaki işlem adımları aşağıdaki alt başlıklarda ayrıntılı olarak bahsedilmiştir.

2.1. MIT-BIH Aritmi Veritabanı

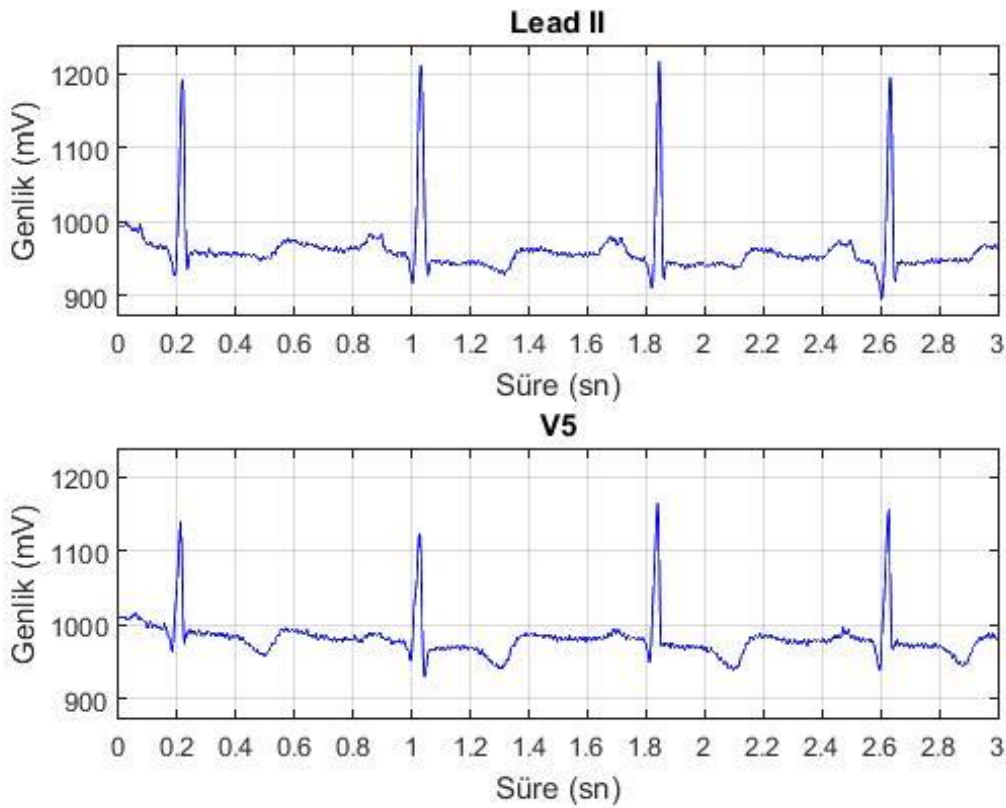
MIT-BIH Aritmi Veritabanı aritmi algılayıcılarının değerlendirilmesi için standart test işaretlerinin koleksiyonu olarak ortaya çıkan ilk veritabanı olma özelliğine sahiptir ve 1980 yılında kullanıma açılmıştır. Kamuya açık hale geldiğinden bu yana tüm dünya üzerinde yaklaşık 500 sitede görülmüştür. Veritabanı, aritmi algılayıcılarının değerlendirmesi dışında kalp dinamiklerinin temel çalışmalarında yardımcı olmuştur. Hatta MIT-BIH Aritmi Veritabanı dünya çapında standart haline gelmiştir.

MIT-BIH Aritmi Veritabanında iki kanallı 48 kayıt bulunmaktadır. Yaklaşık olarak 30dk.'lık EKG verilerini içermektedir. Bu kayıtların tamamı BIH Aritmi laboratuvarı tarafından 1975 ile 1979 yılları arasında çalışılan 47 denek üzerinden oluşturulmuştur. Bunların 25'i erkek ve 22'si kadındır. Erkekler 32-89 yaş aralığında, kadınlar ise 23-89 yaş aralığındadır. 201 ve 202 numaralı kayıtlar tek bir erkek denekten elde edilmiştir.

Kayıtların yaklaşık %60'ı hastanede yatan hastalardan, %40'ı ise ayakta tedavi edilen hastalardan elde edilmiştir. Veritabanı ayakta tedavi edilen hastalardan 24 saatlik 4000 kayıttan rasgele seçilmiş 23 tanesini içermektedir. Bu kayıtlar 100-124 aralığında numaralandırılmıştır (Bazı kayıt numaraları eksiktir). Kayıtların diğer bölümü (200-234 arasında etiketlenmiş 25 kayıt) kayıtlardan rasgele seçilmiş herhangi bir küçük örnek ile

sunulamayacak önemli klinik olayları temsil etmek için aynı EKG setlerinden seçilmiştir [14,37].

Veritabanının ilk bölümü ritim bozukluğu algılayıcılarının rutin klinik uygulamalarda karşılaşılabileceği dalga şekli ve artifaktların geniş bir ölçekte örneklerini içerir. Veritabanının ikinci bölümündeki kayıtlar aritmi algılayıcıları zorlayan, veritabanı kullanıcıları tarafından iyi bilinen ve popüler olmayan çeşitli karmaşık ritim bozukluğu türleridir. Bunlar; ventriküler, supraventriküler, iletim ile ilgili anomaliler, QRS morfolojik ve işaret kalitesi ile ilgili varyasyonlardır.



Şekil 35. 100 numaralı kaydın ilk 3 saniyelik görünümü

Kayıttaki üst işaret çoğu durumda hastanın göğsüne yerleştirilmiş elektrotlar tarafından ölçülen MLII (modified limb lead II) işaretidir. 102 ve 104 numaralı kayıtlarda hastanın ameliyat elbisesi giymesinden dolayı derivasyon II yerine V5 derivasyonu kullanılmıştır. Alt işaret normalde bir V1 kaydıdır. Ancak V2, V5 ve V4 olan çeşitli örnekler de vardır. Şekil 35'te 100 numaralı kayda ait işaretin üç saniyelik kısmı gösterilmektedir. İşaretler dijitale dönüştürülmesi 360 Hz, 10 mV'luk aralık için 11 bit çözünürlükle

yapılmıştır. Her kayıt birbirinden bağımsız konu ile ilgili en az iki kardiyolog tarafından ve vuru-vuru notlar olarak elde edilmiş bilgiler ile desteklenmiştir. MIT-BIH aritmi veri tabanında yaklaşık 109.000 not vardır [14,37].

Bir EKG kaydını belirtmek için bir başlık dosyası (düz metin dosyası), veri dosyası (ikili dosya) ve not dosyası (ikili dosya) kullanılmıştır. Her bir dosyada dosya türüne özel bilgiler tutulur. Başlık dosyasında hastanın yaşı, cinsiyeti ve uygulanan tedavi bilgileri tutulmaktadır. Referans not dosyaları tanımlamaları, özel vuruları, ritimlerini ve ölçülen işaretin kalite karakteristiğini içerir. Veritabanı PhysioBank arşivi web sitesinden ulaşılabilir [211]. Kayıtlar farklı uzantılar ile aynı isimde tutulmaktadır. *atr* uzantısı ile not dosyaları, *dat* ile veri dosyaları ve *hea* ile başlık dosyaları belirtilmektedir [212].

2.2. Ön İşleme

EKG işaretlerinde bulunan gürültü ve dalgalanmalar hesaplanacak öznitelik vektörünü negatif olarak etkilemekte ve işaret işleme algoritmalarının performanslarını düşürmektedir [6]. Farklı hastalardan elde edilen aynı tip EKG işaretlerinde kayda değer değişiklikler gözlenebilmektedir. Hastanın nefes alması, öksürmesi, kıvrılması gibi günlük olağan davranışlarını sergilemesi esnasında alınan kayıtlarda gürültüler sıklıkla bulunur. Birçok biyomedikal işaretin kaynağındaki zayıf seviyeleri nedeniyle kayıttan önce işaretin güçlendirilmesi gerekliliği vardır. İşaret güçlendirilirken doğal olarak gürültüde güçlendirilmiş olur. EKG işaretlerinde bulunan gürültüler bir önceki bölümde ayrıntılı olarak bahsedilmiştir. Bu gürültülerin azaltılması ve işareten çıkarılacak öznitelik vektörlerine olan etkilerinin giderilmesi için işaret normalizasyon ve ön işleme adımlarından geçirilir [24,27,210].

Şekil 35'ten görüldüğü gibi işaretin taban hattı 1000 değeri civarında seyretmektedir. Yapılacak olan matematiksel işlemlerin yükünün azaltılması için ilk olarak işaret ortalaması sıfırlanmaktadır. Bu işlem ile işaretin hastanın karakteristiklerine bağlı özellikleri en aza indirilmiş olur. Sıfır ortalamalı $y(t)$, $t = 1, 2, \dots, L$ işareti eşitlik (63)'teki gibi hesaplanır.

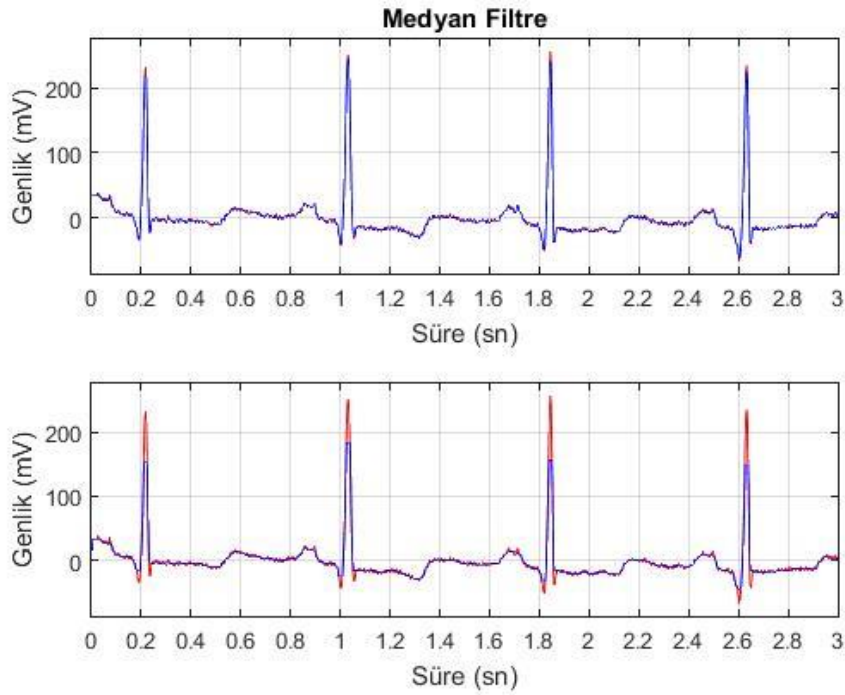
$$y(t) = x(t) - \bar{x} \quad (63)$$

Burada $y(t)$ hesaplanan sıfır ortalamalı işaret, $x(t)$, $t = 1, 2, \dots, L$ ham EKG işareti, \bar{x} değeri $x(t)$ 'nin aritmetik ortalaması ve L işaretin uzunluğudur.

Bir sonraki aşamada pencere genişliği 3 olan bir medyan filtresi kullanılmıştır. Medyan filtresi işarettaki keskin kenarları koruyarak gürültüyü bastıran doğrusal olmayan bir filtredir [213]. Filtrelenmiş işaret $Y(t)$, $t = 1, 2, \dots, L$ eşitlik (64) ile hesaplanır.

$$Y(t) = \text{med} \left\{ \begin{array}{c} y(t-n) \\ y(t-n+1) \\ \vdots \\ y(t-1) \\ y(t) \\ y(t+1) \\ \vdots \\ y(t+n-1) \\ y(t+n) \end{array} \right\} \quad (64)$$

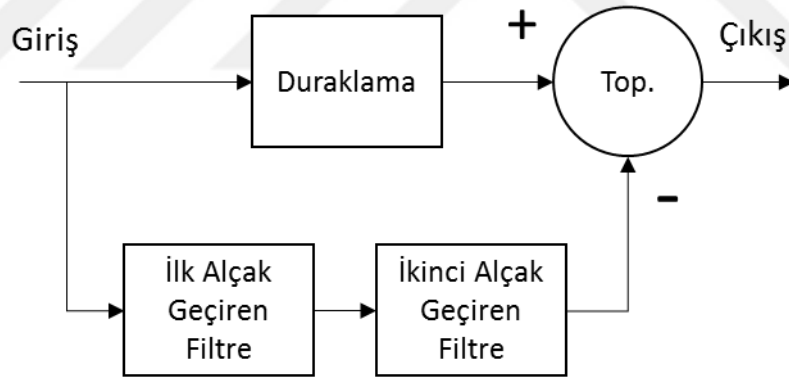
Burada $Y(t)$ filtrelenen işareti, $y(t)$ ise bir önceki adımda hesaplanan giriş işaretini, $2n + 1$ ise pencere genişliğini göstermektedir.



Şekil 36. Medyan Filtre. Üst: pencere genişliği 3 olan filtre sonucu, Alt: pencere genişliği 10 olan filtre sonucu

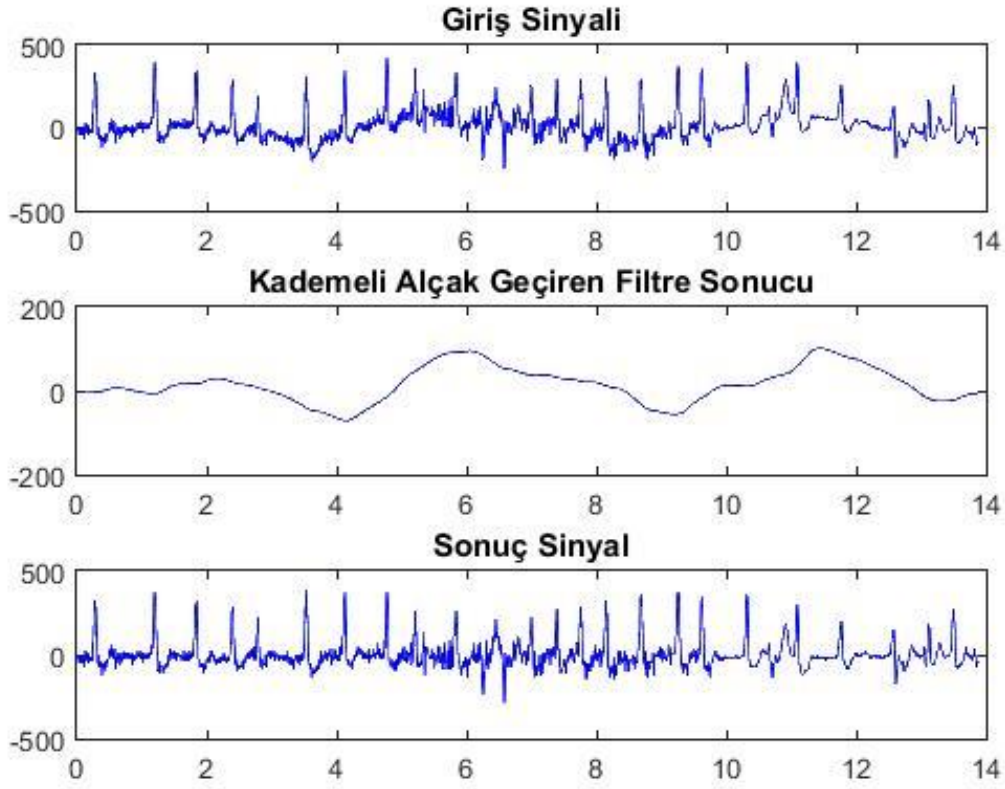
Şekil 36’da pencere genişliği 3 ve 10 olan iki medyan filtre sonucunun ham işaret ile karşılaştırılması görülmektedir. Grafikte mavi renkli çizgi medyan filtre sonucunu, kırmızı renkli çizgi ise giriş işaretini göstermektedir. Şekilden görüleceği gibi pencere genişliği 10 olan medyan filtre doğası gereği işarettaki keskin hatlar olan Q, R ve S noktalarında fazla bozulmaya neden olmaktadır. Bu nedenle çalışmada medyan filtre pencere genişliği 3 olarak seçilmiştir.

Bir sonraki aşamada taban hattından sapmaları elimine etmek için filtre uygulaması yapılmaktadır. Sırasıyla 2 Hz ve 0,5 Hz frekans bileşenlerini kademeli olarak silen bir alçak geçiren filtre eşitlik (64)’te hesaplanan $Y(t)$ işaretine uygulanır. Taban hattından sapmalara sebep olan gürültünün frekans bileşenleri genellikle 0,5 Hz’in altındadır [214]. Ancak bazı durumlarda bu daha yüksek olabilir. Bu nedenle frekans limiti 2 Hz olarak belirlenmiştir. Alçak geçiren filtrelerden elde edilmiş sonuç taban hattından sapmadır. Bu işaret ham işaretten çıkarılarak taban hattı sapması giderilmiş olur. Şekil 37’de taban hattından sapmaları ortadan kaldıran filtrenin genel şeması gösterilmiştir.



Şekil 37. Kademeli alçak geçiren filtre uygulamasının blok diyagramı

Şekil 38’de giriş işaret, taban hattında sapma ve sonuç işaret gösterilmiştir. Bu örnekte MIT-BIH Aritmi Veritabanındaki 203 numaralı kaydın ilk 14 saniyelik kısmı kullanılmıştır.



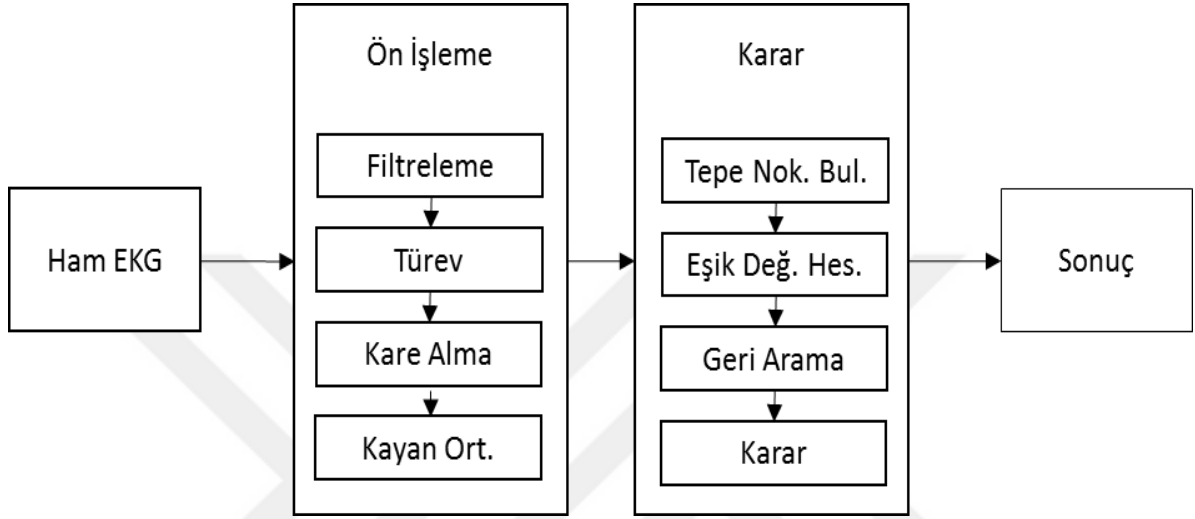
Şekil 38. Taban hattı sapmasının filtrelenmesi.

2.3. QRS Tespiti

EKG işaretlerinin işlenmesinde kullanılan uygulamalar için yapılması gereken en önemli görevlerden bir tanesi QRS dalgasının, özellikle R tepesinin doğru olarak algılanmasıdır. QRS kompleksine ek olarak EKG'de P ve T dalgaları, güç hattı sapmaları, EMG gürültüsü, hastanın hareket etmesinden kaynaklanan gürültüler bulunmaktadır. QRS kompleksini diğer gürültü ve dalgalardan ayırt etmek önemlidir.

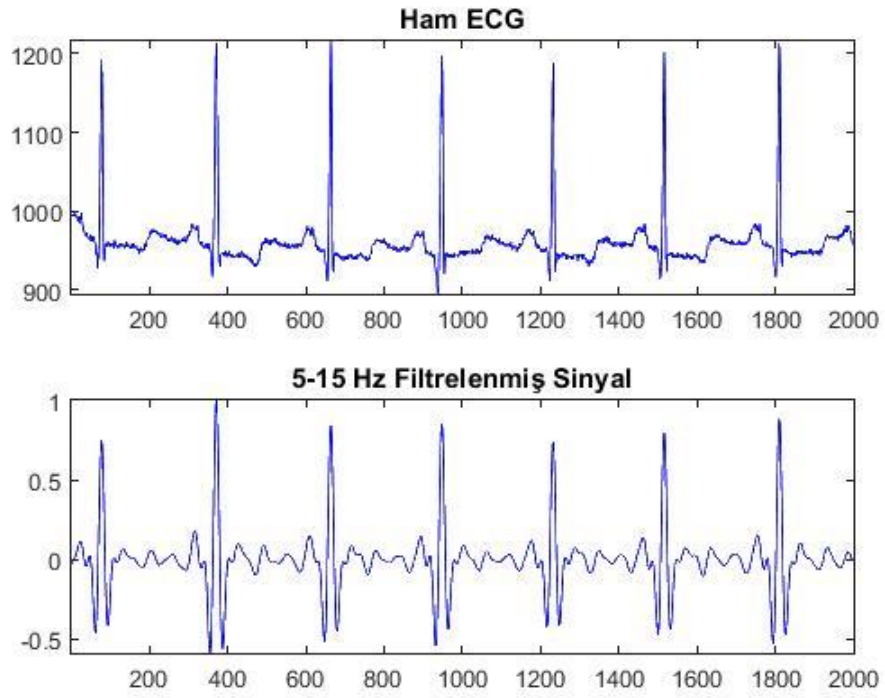
Bu çalışmada Pan ve Tompkins [119] tarafından geliştirilen QRS algılama algoritması kullanılmıştır. QRS kompleksinin enerjisi yaklaşık olarak 5-15 Hz arasında olduğu bilinmektedir. Bu algoritma R tepesinin analiz edilmesi temeli üzerine kurulmuştur. Algoritmada işareti sadece istenen frekans aralığına indirmek için alçak ve yüksek geçiren filtreler kullanılmıştır. Bu sayede istenmeyen gürültüler, P ve T dalgaları elimine edilmektedir.

QRS algılama algoritması giriş olarak ham EKG işaretini alır. Bir önceki başlıkta anlatılan filtreleme işlemleri sonucunda elde edilmiş işaret, QRS kompleksi belirlendikten sonra vuru ayrıştırma adımında giriş olarak kullanılacaktır. Pan ve Tompkins algoritmasının akış diyagramı Şekil 39’da verilmiştir.

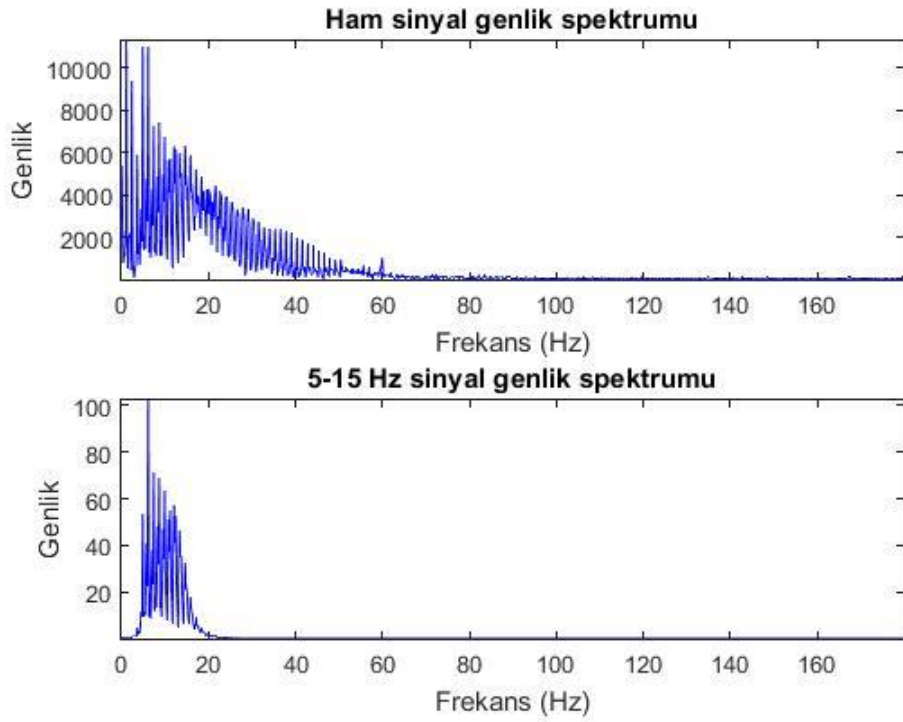


Şekil 39. QRS Algılama algoritmasının genel şeması.

Filtreleme aşamasında ilk olarak işaret 5-15 Hz bandında olacak şekilde filtrelenir. Tompkins’in çalışmasında bant geçiren filtre alçak geçiren ve yüksek geçiren filtrenin birlikte kullanılması ile gerçekleştirilmiştir. Alçak geçiren filtre kesme frekansı 15 Hz, yüksek geçiren filtre kesme frekansı olarak 5 Hz seçilmiştir. Bu filtreleme işlemi sonucunda elde edilen işaret karşılaştırmalı olarak Şekil 40’ta gösterilmiştir. Şekil 41’de ise ham EKG ve filtrelenen işaret sonucunun frekans spektrumu gösterilmektedir.

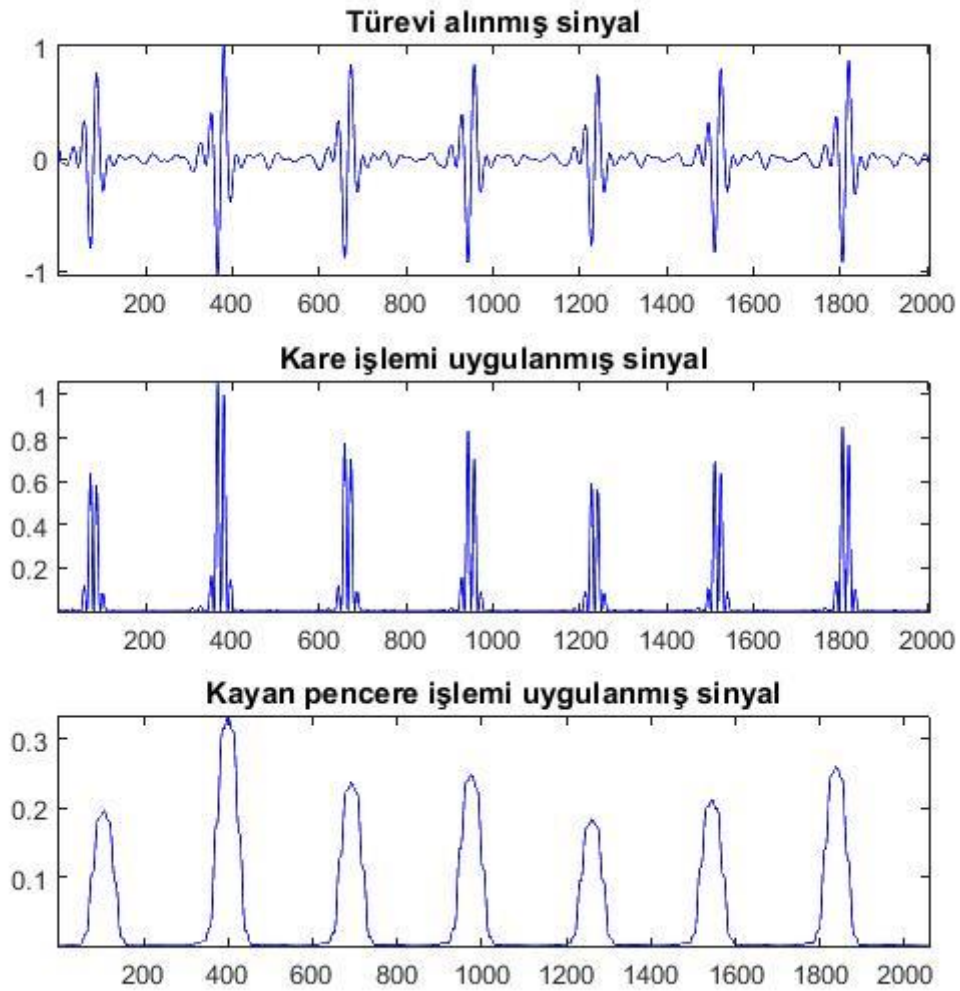


Şekil 40. Orijinal EKG ve 5-15 Hz bandında filtrelenmiş sonucu



Şekil 41. Ham EKG ve Filtrelenen işaretin genlik spektrumu

Bir sonraki aşamada gerçekleştirilecek olan türev operatörü R noktasını daha belirgin hale getirmek için kullanılır. Bu süreç QRS kompleksi yüksek frekans özelliklerini güçlendirir ve P ve T dalgalarının düşük frekanslarını zayıflatır. Devam eden aşamada ise kare işlemi gerçekleştirilir. Bu süreç türev işleminin çıktısının yoğunluğunu arttırmaktadır. Son aşama ise kayan pencere toplamı olarak adlandırılan zaman ortalama sürecidir. Bu toplama operatörü 150 ms artımla adım adım zaman domeninde taşınarak karesi alınmış işaretin dalga şeklinin altında kalan alan toplamını hesaplamaktadır. Şekil 42’de bu üç aşama sonucunda elde edilen çıktılar gösterilmiştir.



Şekil 42. Türev, Kare İşlemi ve Kayan Pencere işlemleri

Karar aşaması QRS algılama algoritmasının son aşamasıdır. R tepesinin yeri algoritma tarafından tespit edilir. Son aşama gerçekleştirilen kayan pencere sürecinin çıktısı yüksek genlikli ve bazı küçük genlikli tepelerden oluşmaktadır. Algoritma burada küçük tepelerin R tepesimi yoksa gürültü veya genliği yüksek bir T tepesimi olduğunu anlamaya çalışır.

Bu aşamada belirlenen bir aralıkta bulunan tüm tepelerin (yerel maksimumlar) yeri tespit edilir ve tespit edilen yerlerin R tepesi, T tepesi veya gürültü olup olmadığı belirlenir. Kayan pencere sürecinin çıktısı analiz edilirken algoritma EKG işaretinin kalitesine göre kendini sürekli olarak güncelleyen iki set eşik değeri kullanır: SIN_SEV, SIN_ESK ve GUR_SEV, GUR_ESK. SIN_ESK ve GUR_ESK değerlerine ilk değerleri 2 saniyelik kısa eğitim aşaması ile atanmaktadır. İlk olarak tüm tepeler değerlendirilmesinde (TEPEI) işaret veya gürültü olarak sınıflandırmak için bu eşik değerleri kullanılır.

Eğer $TEPEI > SIN_ESK$ ise bu konum aday QRS olarak belirlenir ve işaret seviyesi güncellenir.

$$SIN_SEV = 0,125 \times TEPEI + 0,875 \times SIN_SEV \quad (65)$$

Eğer $GUR_ESK < TEPEI < SIN_ESK$ ise bu konum gürültü tepesi olarak belirlenir ve gürültü seviyesi güncellenir.

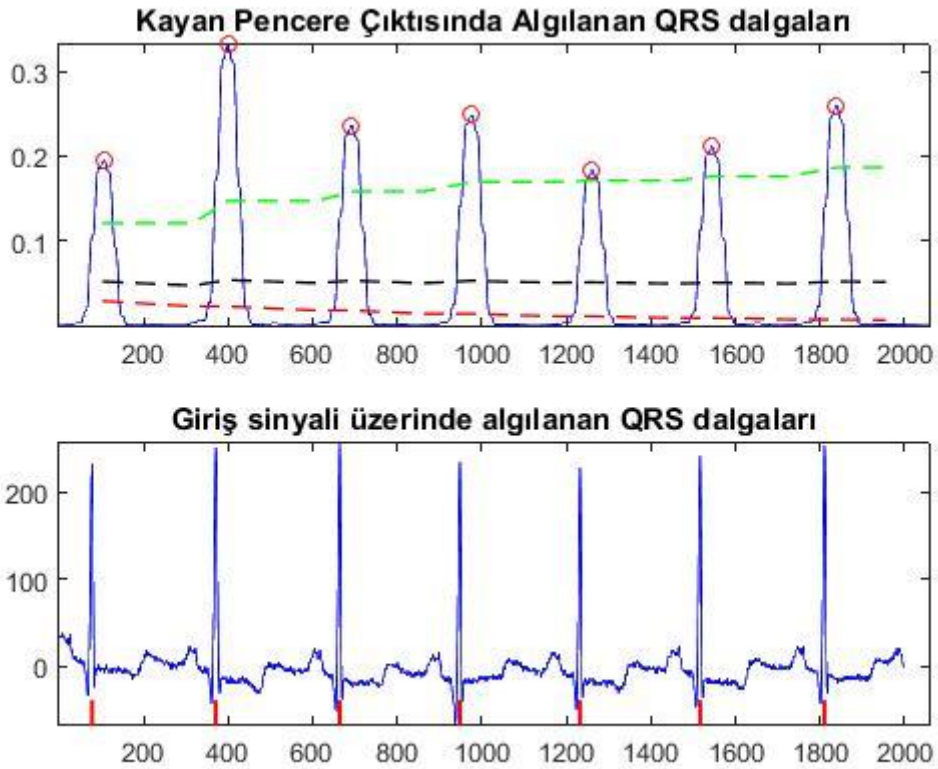
$$GUR_SEV = 0,125 \times TEPEI + 0,875 \times GUR_SEV \quad (66)$$

İşaret seviyesi ve gürültü seviyesi değerlerinin bu yeni yaklaşımlarına göre bu noktadaki EKG için işaret ve gürültü eşik değerleri eşitlik (67)'deki değerlere güncellenir.

$$\begin{aligned} SIG_ESK &= SIN_SEV + 0,25 \times (SIN_SEV - GUR_SEV) \\ GUR_ESK &= 0,5 \times SIG_ESK \end{aligned} \quad (67)$$

Bu güncellemeler ile zayıf kalitede olarak varsayılan işaret bölümlerinde eşik değerleri kademeli olarak düşürülür.

Algoritmada bu aşamadan sonra kaçırılacak QRS kompleksleri için geri arama yapılmaktadır. Geri arama prosedürünün tetiklenmesi için şu anda algılanan son R-R aralığı ölçülür ve 1,66 katından fazla aralık olmuşsa geri arama gerçekleştirilir ve kaçırılan tepenin iki eşik değeri arasında olduğu düşünülür. Bu durumda R-R aralıklarından faydalanılır ve R-R aralıklarının ölçülmesinde iki değer kullanılır. Bir tanesi son 8 R-R aralığının ortalamasıdır. Diğeri ise son R-R aralığıdır. R-R aralığının düzenli olmadığı durumlarda algılamayı geliştirmek için eşik değeri düşürülmüştür.



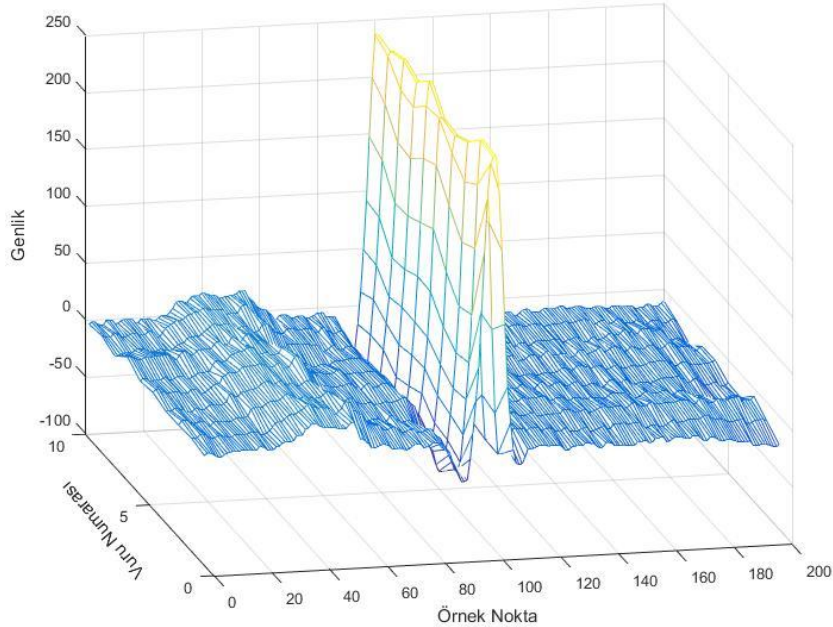
Şekil 43. QRS algılama işlemi sonucu

Bir diğer düzenleme operatörü ise aynı periyottaki çoklu algılamaların elenmesidir. Bir QRS algılandıktan 200ms sonrasında kadar yeni bir QRS fizyolojik olarak olması imkansızdır. Algoritma bu fizyolojik kısıtlamadan dolayı QRS adaylarını eler. QRS olarak algılanabilecek T dalgalarının elemesi gerçekleştirilir. Eğer QRS adayı 200ms periyodu sonunda ancak önceki QRS'nin 360ms süresi içerisinde ise bu tepenin bir sonraki yeni QRS ya da anormal olarak belirgin bir T dalgası mı olduğunu belirlenir. Karar bu noktadaki dalga biçiminin ortalama eğimine dayanır. Buradaki eğim bir önceki QRS kompleksinin

eğiminin yarısından az ise T dalgasının yavaş değişen davranışı olarak algılanır, aksi takdirde QRS algılaması doğrudur. Bu son aşama ile birlikte QRS algılama tamamlanmış olur. Şekil 43’de nihai QRS algılama sonucu gösterilmiştir. Üst şekilde kayan pencere sonucunda algılama işlemi verilmiştir. Burada kırmızı daireler algılanan tepe noktalarını, kırmızı kesikli çizgi gürültü seviyesini, yeşil kesikli çizgi işaret seviyesini ve siyah kesikli çizgi de eşik seviyesini göstermektedir. Alt şekilde ise giriş işareti üzerinde kırmızı dikey çizgiler ile algılanan QRS noktaları gösterilmiştir.

2.4. Vuru Ayırıştırma

QRS algılaması gerçekleştirildikten sonra filtrelenmiş işaret üzerinde vuru ayırıştırma işlemi yapılır. Her bir vuru için 200 (556ms) örnek noktadan oluşan pencere genişliği belirlenmiştir. Burada R tepesi 100. noktayı temsil edecek şekilde yerleştirilmiş, 99 öncesi ve 100 sonrası alınarak bir kalp vurusunu temsil eden vuru işareti elde edilmiştir. Şekil 44’de MIT-BIH Aritmi Veritabanındaki 100 numaralı kaydın ilk 10 vurusu ayrııştırılarak oluşturulmuş verinin 3 boyutlu görünümü gösterilmiştir.



Şekil 44. 100 numaralı kaydın ilk 10 vurusu

Elde edilmiş vurular ilerleyen başlıklarda öznitelik çıkarma ve sınıflandırma amacı ile kullanılacaktır.

2.5. Erken Karıncık Kasılması Vurusunun Algılanması

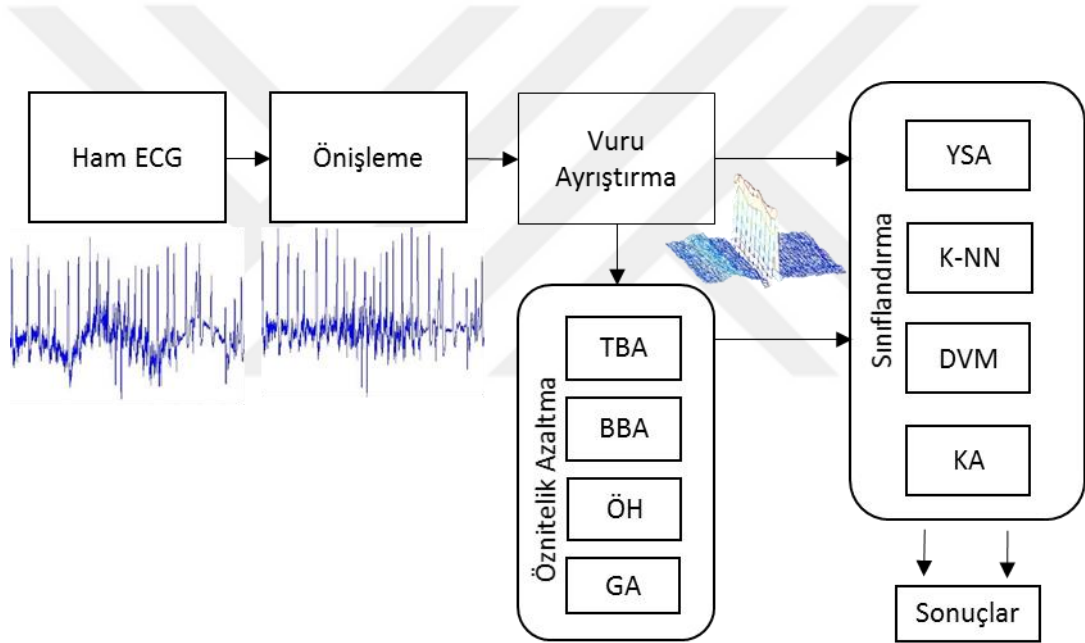
Kardiyak aritmiler kalp hastalıklarının en önemli habercileridir ve anormal kalp ritimleridir. Aritmiler kalbin elektriksel sistemindeki iletimde, kalp atım oranında ve kalp atım devamlılığında olabilecek değişiklikler sebebiyle oluşur [13]. Çeşitli aritmiler arasında EKK en önemlilerinden bir tanesidir [215]. EKK olması gereken zamandan önce meydana gelen kalp atışlarıdır. EKK karıncık kasının herhangi bir bölgesindeki bir odaktan doğabilecek bir tetikleme ile karıncık kaslarının erken depolarizasyonu sonucu ortaya çıkan ve daha çok yetişkinlerde görülen yaygın bir aritmi türüdür. EKG işareti üzerinde incelendiğinde erken oluşmuş bir QRS kompleksi, takip eden uzun düzenleyici bir duraklama EKK'nin karakteristik özelliğidir. EKK'nin diğer özelliği QRS kompleksinden önce P dalgasının olmayışdır. EKK QRS kompleksinin daha geniş olmasını sağlayan bozulmalara sebep olur [215,216].

EKK genel nüfusun %1 ile %4'ü arasında rastlandığı tahmin edilen yaygın bir aritmidir ve genellikle yaşa bağlı olarak değişmektedir. 11 yaş altı çocuklarda görülme sıklığı %1 iken 75 yaş üstü bireylerde görülme sıklığı %69'dur. EKK'lar normal sağlıklı nüfusta genellikle 24 veya 48 saatlik Holter kayıtlarının incelenmesi sonucu algılanabilir. Genellikle yapısal kalp hastalıkları ile birlikte görülür ve ani ölüm riskini arttırır. Bireyde bilinen bir kalp rahatsızlığı olmamasına rağmen EKK görülme ihtimali vardır. Değerlendirilmesi ve tedavisi karmaşık bir süreçtir [210,215–217].

Çalışmanın bu bölümünde EKK aritmilerinin sınıflandırılması için etkili bir yöntem önerilmiştir. Çalışmanın öncelikli amacı EKK aritmisinin sınıflandırılmasında sınıflandırma doğruluğunun arttırılmasıdır. Bu amaçla geliştirilecek olan sistemlerde işlem yükünün az olması için herhangi bir şekilsel veya matematiksel öznitelik çıkarma aşaması kullanılmamıştır. İşaretin zaman serisi (genliği) sınıflandırma algoritmalarında giriş vektörü olarak kullanılmış ve algoritmaların performansları değerlendirilmiştir. Karşılaştırma için zaman serilerini giriş verisinin boyutu BBA, TBA ve ÖH kullanılarak azaltılmış ve sonuçlar irdelenmiştir. Bunlara ek olarak GA ile seçilen öznitelikler kullanılarak boyut indirgenmiş

ve sonuçlar karşılaştırılmıştır. Deneysel sonuçlar için YSA, K-NN, DVM ve KA algoritmaları farklı sınıflandırma şemalarında kullanılmıştır [210].

EKG veri sınıflandırma karar süreci diğer sınıflandırma süreçlerine benzer şekilde 3 aşamadan oluşmaktadır; Ön işleme, öznelik çıkarımı ve sınıflandırma aşaması. Daha önce yapılan çalışmalarda ön işleme aşamasında benzer yöntemler kullanılmıştır. Araştırmacılar öznelik çıkarımı ve sınıflandırma süreçlerinde farklı modellere odaklanmıştır. Şekil 45’de EKK virusunun sınıflandırılması için tasarlanan sistemin şeması gösterilmiştir. Şekil 31’deki vuru sınıflandırma şemasına benzer olarak EKK virusunun sınıflandırılması için tasarlanan sistem alt bloklar halinde verilmiş ve kullanılan yöntemler belirtilmiştir.



Şekil 45. EKK virusunun sınıflandırılmasını gösteren şema

2.5.1. Erken Karıncık Kasılması Virusunun Algılanması İçin Veri Seti

Ön işleme ve vuru ayrıştırma işlemleri bir önceki aşamada ayrıntılı olarak verilmişti. Ayrıştırılmış vurulardan testlerde kullanılacak test verileri Tablo 5’te gösterilmiştir.

Tablo 5. EKK sınıflandırma için veri setindeki vuruların dağılımı

<i>Kayıt No</i>	<i>Vuru</i>		<i>Toplam</i>
	<i>Normal</i>	<i>EKK</i>	
<i>100-105</i>	401	42	443
<i>106-109</i>	99	622	721
<i>111-115</i>	300	1	301
<i>116-119</i>	300	568	868
<i>121-124</i>	300	51	351
<i>200-203</i>	400	456	856
<i>205-209</i>	300	296	596
<i>210-214</i>	300	669	969
<i>215-220</i>	400	389	789
<i>221-228</i>	400	400	800
<i>230-234</i>	300	6	306
<i>Toplam</i>	3500	3500	7000

Çalışmada MIT-BIH aritmi veritabanında bulunan ve II derivasyonu içeren kayıtlardan normal ve EKK vurularını içeren tüm kayıtlar kullanılmıştır. Normal vuru hemen hemen her kayıttan olduğu için her kayıttan seçilmiştir. EKK vuruları kayıtlarda düzensiz olarak dağıldığı için farklı sayılarda seçilmiştir. Veri tabanında 7000 civarı EKK vurusu vardır ve bunlardan çalışmada 3500 tanesi deneysel sonuçlar için kullanılmıştır. Her kayıttan yaklaşık 100 tane seçilerek 3500 normal vuru çalışmaya dahil edilmiştir [210].

EKK vurularının sınıflandırılması için herhangi bir öznitelik çıkarma yöntemi kullanılmamış, işaretin bir vuruluk zaman serisi (200 örnek) sınıflandırıcılar için giriş verisi olarak kullanılmıştır. Sistem performansının artırılması için bu giriş verisinin boyutu öznitelik azaltma ve öznitelik seçim yöntemleri kullanılarak azaltılmış ve sonuçları karşılaştırılmıştır. TBA, BBA ve ÖH öznitelik azaltma yöntemleri ve GA öznitelik seçim yöntemi kullanılmıştır. Giriş verisinin 200 örnekten oluşmasından dolayı sıralı ileri arama ve sıralı geri arama yöntemleri kullanılmamıştır.

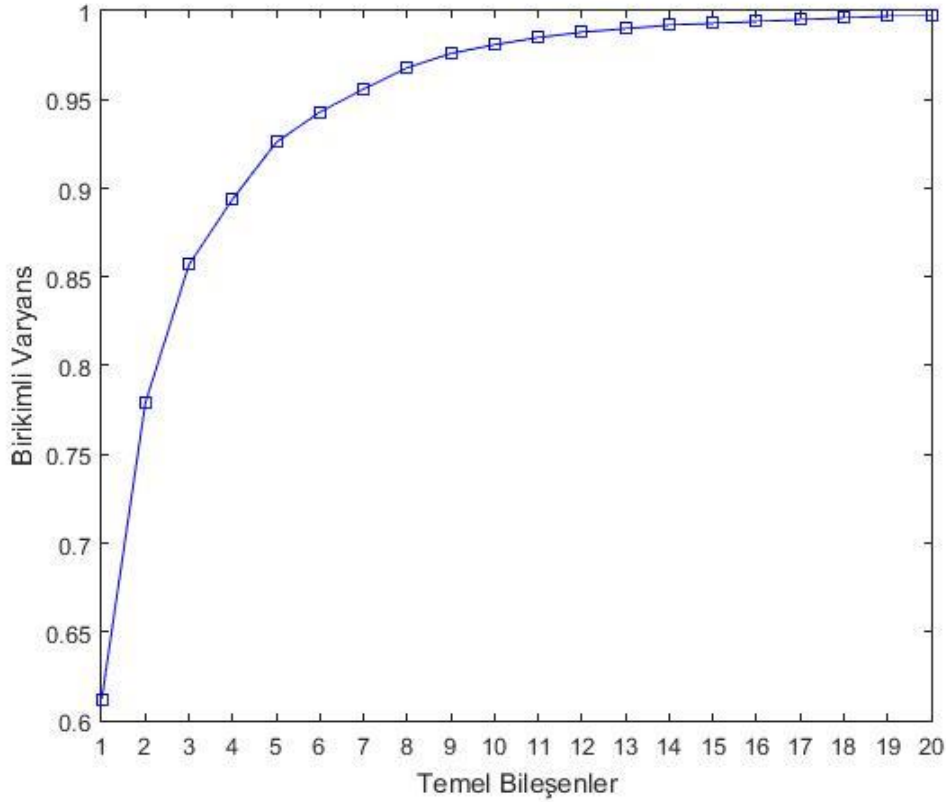
Bu algoritmalar kullanılarak sınıflandırma algoritmalarının giriş vektör boyutları azaltılmış, sınıflandırma başarımında herhangi bir azalma olmadan, hatta artış sağlanarak sistemin karmaşıklığı azaltılmıştır. Tek vuruluk zaman serisi ve boyutu indirgenmiş veriler sınıflandırmada karşılaştırmalı olarak kullanılmış ve algoritmalarının eğitiminde ve testinde kayda değer hızlanma sağlanmıştır.

2.5.2. EKK Sınıflandırma İçin Boyut İndirgeme

Önceki bölümde ayrıntılarıyla bahsedilen TBA, muhtemel ilişkili gözlem kümesi içinden temel bileşenler olarak adlandırılan birbiri ile ilişkisiz öznelik kümesine dönüşüm için dik dönüşümü kullanan sayısal bir yöntemdir [164]. BBA ise gözlemlenen rasgele veriyi birbirinden bağımsız bileşenlere dönüştüren çok yönlü istatistiksel bir yöntemdir [165]. ÖH, Kohonen tarafından geliştirilmiş, danışmansız öğrenme yöntemini kullanan bir yapay sinir ağıdır. ÖH yüksek boyutlu giriş uzayını daha düşük boyuta (özellikle iki boyut) etkili ve kolay bir şekilde haritalayan bir yaklaşımdır. Bu yöntemlerden farklı olarak GA, 1975 yılınca John Holland tarafından geliştirilmiş [159], çalışmada öznelik seçimi için kullanılmıştır.

Sınıflandırma aşamasına geçilmeden önce veriyi en iyi temsil edecek temel bileşenlerin sayısını belirlemek için ızgara arama yöntemi kullanılmıştır. Bu arama işleminde bir sınıflandırıcı değerlendirme fonksiyonu olarak kullanılarak her adımda parametre değerleri hesaplanmaktadır. Bu amaçla 1-30 arasındaki temel bileşenler hesaplanarak her adımda sonuçlar değerlendirilmiş ve en iyi sonucu verecek temel bileşen sayısı hesaplanmaya çalışılmıştır [210].

Şekil 46'da hesaplanan ilk 20 temel bileşene ait birikimli varyans değerleri gösterilmiştir. Şekilden görüldüğü gibi temel bileşenler hesaplanırken birikimli varyans küçük değerlerden başlayarak temel bileşen sayısı arttıkça hızlı bir şekilde yükselmiştir. Temel bileşen sayısı 5 iken 0,926 olan birikimli varyans temel bileşen sayısı 20 olduğunda 0,997 değerine ulaşmıştır. Temel bileşenlerin değerlendirilmesinde 1-30 aralığındaki temel bileşenler sınıflandırma sonucu için test edilmiş ve en iyi sonucu veren temel bileşen sayısı olarak 17 değeri bulunmuştur. Temel bileşen sayısı 17 olduğunda birikimli varyans değeri 0,995 olarak hesaplanmıştır. Bu nedenle sınıflandırma aşamasında temel bileşenler kullanılırken temel bileşen sayısı 17 olarak alınmıştır.



Şekil 46. Temel bileşenlerin birikimli varyans değerleri

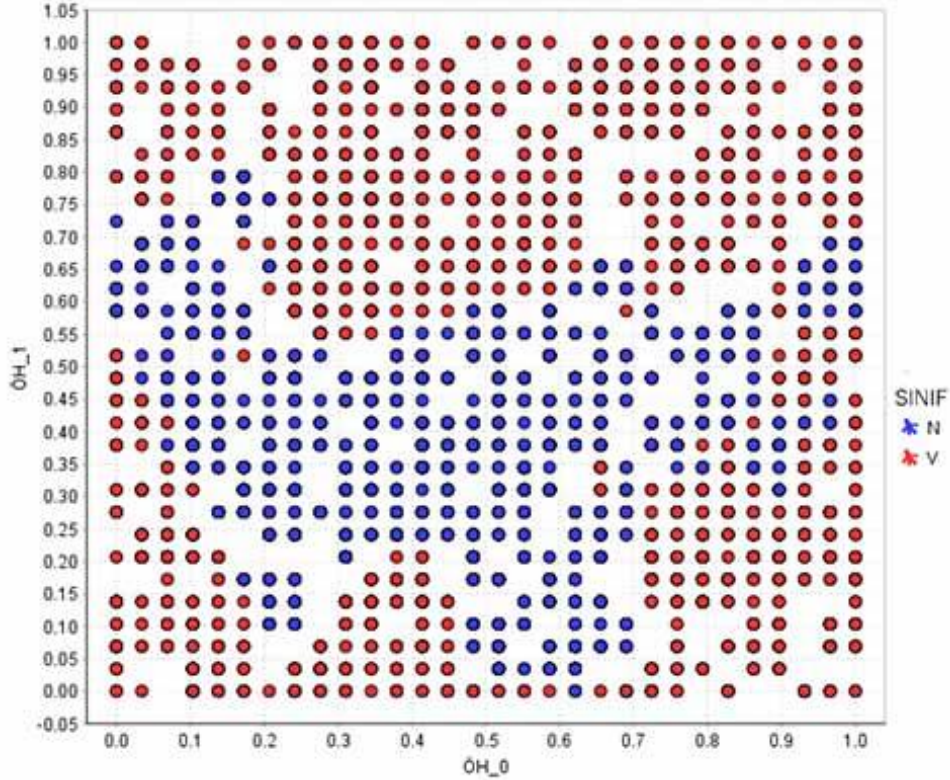
Çalışmada diğer öznelik setleri ile karşılaştırmada kullanılmak üzere bir vuruluk işaretin zaman serisinin bağımsız bileşenleri hesaplanmıştır. Bağımsız bileşenleri hesaplamak için hızlı BBA algoritması [165] kullanılmıştır. TBA'daki arama yapısına benzer şekilde 1-30 arasındaki tüm bağımsız bileşenler, en iyi sınıflandırma sonucunu verecek bağımsız bileşen sayısını bulmak için parametre optimizasyonu yapılmıştır. Deneyle sonuçunda bağımsız bileşen sayısı 10 olarak hesaplanmıştır.

Çalışmada kullanılan bir diğer boyut indirgeme yöntemi ise ÖH'dir. ÖH kullanılarak veri setinin boyutu iki boyutlu uzaya indirgenmiştir. Çalışmada ÖH çıktısı iki ve üç olarak testler gerçekleştirilmiş ve iki seçildiğinde daha iyi sonuçlar elde edilmiştir [210]. ÖH kullanılarak

- Ağın boyutu 30,
- Öğrenme tur sayısı olarak 30,
- Başlangıçtaki öğrenme katsayısı 0,8,

- Bitiş öğrenme katsayısı 0,01,
- Başlangıçtaki (ilk turdaki) daire çapı 10,
- Son tur daire çapı 1

olarak belirlenmiştir. Veri setinin ÖH ile iki boyutlu uzaya dönüştürülmüş hali Şekil 47’de gösterilmiştir.

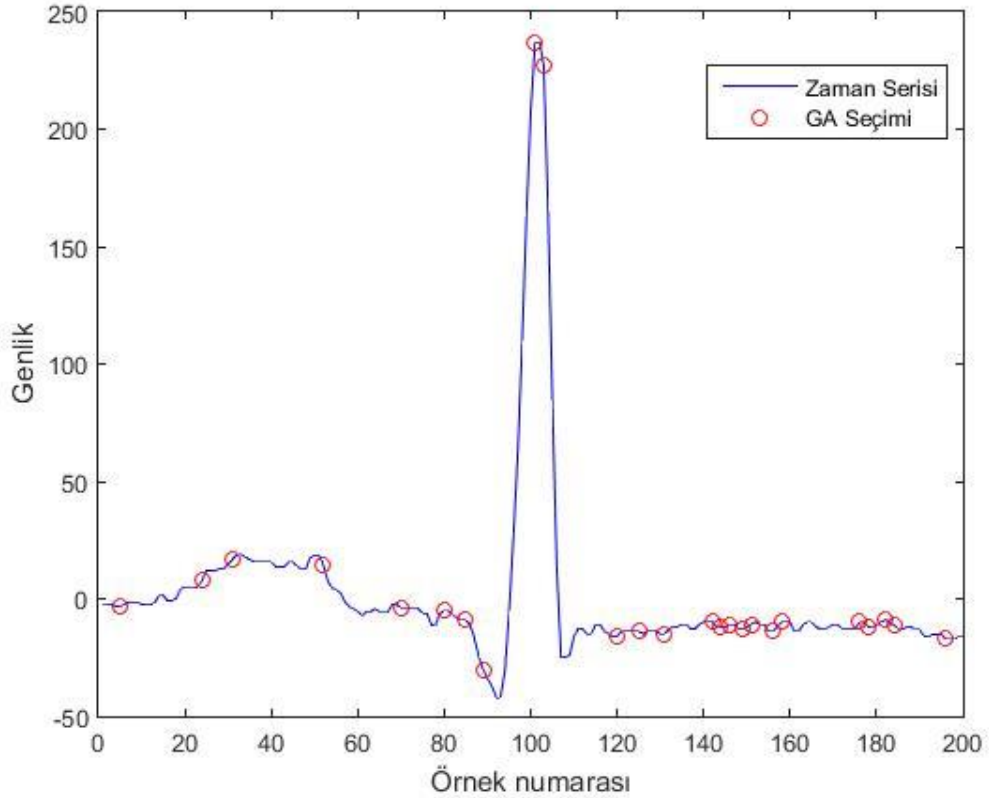


Şekil 47. ÖH ile boyutu indirgenmiş test verisi

2.5.3. Genetik Algoritma ile Öznitelik Seçimi

Yukarıdaki yöntemler 200 örneklilik giriş verisi üzerinden matematiksel yöntemler kullanarak daha düşük boyutta yenilenmiş öznitelikler üretmektedir. Bu yöntemlerden farklı olarak 200 örneklilik giriş verilerinden veriyi temsil edecek en iyilerinin seçilmesi farklı bir yaklaşımdır. Bu amaçla GA 200 noktadan oluşan giriş verisinin boyutunu azaltmak için kullanılmıştır. Bu çalışmada turnuva temelli seçim stratejisi kullanan standart bir GA kullanılmıştır. Testlerde kullanılan GA parametreleri aşağıda belirtilmiştir;

- Popülasyon boyutu: 5
- Maksimum jenerasyon sayısı: 30
- Çaprazlama olasılığı: 0,5
- Mutasyon Olasılığı: 1/200 (200 giriş vektör boyutu)
- En az seçilebilecek öznitelik sayısı: 15
- En fazla seçilebilecek öznitelik sayısı: 25 – 50 – 100



Şekil 48. Bir vuruluk işaretin zaman serisi ve GA ile seçilmiş örnekleri

Öznitelik sayısının büyük olması, algoritmadaki döngü sayısını hesaplanamayacak düzeyde artırdığından dolayı bu şekilde seçim stratejisi olan algoritmalarda uygun olmamaktadır. Bu nedenle öznitelik seçim işlemi, seçilebilecek maksimum öznitelik sayısını belli sayılarda limitleyerek GA test edilmiştir. 25, 50 ve 100 değerleri seçilebilecek en fazla öznitelik sayısına sınır olarak belirlenmiş ve sonuçları irdelenmiştir. Bu sınırlar için sırasıyla 25, 50 ve 95 adet öznitelik GA tarafından seçilmiştir. Şekil 48’de üst limit olarak 25 belirtildiğinde GA tarafından seçilmiş zaman serisi noktaları gösterilmiştir [24].

2.5.4. EKK Vurularının Sınıflandırılması

Bir vuruluk EKG işaretinin 200 noktadan oluşan zaman serisi ve BBA, TBA, ÖH ve GA ile boyutu indirgenmiş giriş verileri sınıflandırma algoritmaları tarafından sınıflandırılmış ve sonuçları karşılaştırmalı olarak irdelenmiştir. Her sınıflandırma şeması için sınıflandırma başarımını arttıracak parametre optimizasyon işlemleri gerçekleştirilmiştir.

Tablo 6. YSA gizli katman düğüm sayısı ve farklı giriş vektörlerinin başarıma etkisi

Gizli katman düğüm sayısı	Zaman Serisi	BBA	TBA	ÖH	GA
2	0,9729	0,961	0,9677	0,7879	0,9734
4	0,9783	0,9704	0,9791	0,8744	0,9797
6	0,9873	0,9766	0,9829	0,924	0,9809
8	0,983	0,9771	0,9827	0,9343	0,9819
10	0,9846	0,9779	0,9814	0,9466	0,9817
12	0,98	0,9773	0,982	0,9574	0,9807
14	0,981	0,977	0,9864	0,9703	0,9817
16	0,982	0,9776	0,9831	0,9756	0,9794
18	0,9801	0,9783	0,9817	0,9723	0,9799
20	0,9809	0,9779	0,9851	0,977	0,9794

Testlerde ilk olarak çok katmanlı YSA mimarisi kullanılmıştır. Giriş katmanı belirtilen giriş vektörünün boyutunda, çıkış katmanı ise sınıf sayısı ile aynı olacak şekilde yani iki olarak belirlenmiştir. Tasarlanan YSA mimarisinde bir adet gizli katman olup gizli katmanın düğüm sayısı 10 olarak belirlenmiştir. Testlerde daha fazla gizli katman sayıları denenmiş ancak tek gizli katman kullanıldığında daha iyi sonuçlar elde edilmiştir. Bu nedenle YSA'da bir gizli katman kullanılmıştır. Gizli katmandaki düğüm sayısı deneysel sonuçlar ile elde edilmiştir. Deneylerde 2-20 arasındaki sayılar gizli katman düğüm sayısı olarak belirlenmiş ve elde ettikleri sınıflandırma başarımları incelenmiştir. Tablo 6'da giriş vektörü olarak Zaman serisi, TBA, BBA, ÖH ve GA ile elde edilmiş öznitelikleri kullanılarak YSA'nın farklı gizli katman düğüm sayıları ile elde edilmiş test sonuçları gösterilmiştir.

Çalışmada kullanılan YSA geri yayılım algoritması kullanarak eğitilmiştir. Eğitim aşamasında tekrar sayısı parametresi 500 ve öğrenme katsayısı 0,3 olarak belirlenmiştir. Hata eşiği parametresi ise 0,00001 olarak ayarlanmıştır.

Çalışmada kullanılan bir diğer sınıflandırıcı K-NN'dir. K-NN algoritmasında en iyi K değerini hesaplayabilmek için ızgara arama yöntemi kullanılmıştır. Bu işlem sonucunda K değeri bir olarak bulunmuştur ancak test aralığındaki tüm K değerleri kayda değer yüksek başarılar elde etmiştir. Sonuçları elde etmek için öklit uzaklık fonksiyonu kullanılmıştır. K-NN sınıflandırıcının sınıflandırmadaki başarımından dolayı öznitelik azaltma aşamasında değerlendirme fonksiyonu olarak K-NN sınıflandırıcının başarımı kullanılmıştır.

DVM test sonuçlarını elde etmek için K-NN aşamasındaki aramaya benzer olarak ızgara arama ile parametre optimizasyonu gerçekleştirilmiştir. Parametre optimizasyonu için çekirdek fonksiyonları, karmaşık sabit DVM parametresi C ve Gamma parametreleri en iyi sonucu verecek şekilde optimize edilmeye çalışılmıştır. Aramada çekirdek fonksiyonları için polinom, RBF, sigmoid ve linear olmak üzere 4 çekirdek fonksiyonu denenmiştir. C parametresi için 0-1000 aralığında üstel olarak artan 12 farklı parametre denenmiştir. Gamma parametresi olarak 0-100 aralığında üstel olarak değişen 18 farklı Gamma değeri test edilmiştir. Optimizasyon adımı sonucunda çekirdek fonksiyonu olarak polinom fonksiyonu belirlenmiş, C = 100 ve Gamma = 4 olarak hesaplanmıştır.

Tüm sınıflandırıcılar giriş verisi olarak zaman serilerinin yanı sıra önceki adımda ayrıntıları açıklanan boyutu azaltılmış giriş vektörlerini kullanılmıştır. Sınıflandırma aşamasına geçilmeden önce en iyi sonucu verecek parametrelerin hesaplanması için parametre optimizasyonu işlemleri gerçekleştirilmiştir. Sınıflandırma işlemlerinde 10-kere çapraz doğrulama yöntemi kullanılmıştır ve veri setindeki 10 alt kümeden her biri bir kez tahmin edilmiş olmaktadır.

2.5.5. EKK Vurularının Sınıflandırma Sonuçları

Sınıflandırma algoritmalarının tanıma başarımlarının ölçümü için *doğruluk*, *duyarlılık* ve *özgünlük* değerleri hesaplanmıştır. Tablo 7 öznitelik vektörü olarak zaman serisi kullanan sınıflandırıcıların sınıflandırma performans parametrelerini göstermektedir. Sınıflandırıcılar giriş verisi olarak zaman serisini kullandıklarında K-NN sınıflandırıcısı %99,56 doğruluk değeri ile en yüksek başarıyı elde etmiştir.

Tablo 7. Giriş verisi olarak zaman serisi için sınıflandırma sonuçları (%)

	Doğruluk	Duyarlılık	Özgünlük
<i>K-NN</i>	99.56	99.29	99.83
<i>YSA</i>	98.46	99.06	97.86
<i>DVM</i>	98.09	96.86	99.31
<i>KA</i>	97.96	97.40	98.51

Tablo 8-11 TBA, BBA ve ÖH ile hesaplanan ve GA ile seçilen öznitelik vektörlerini giriş vektörü olarak kullanan sınıflandırıcıların başarımları karşılaştırmalı olarak gösterilmiştir. Sınıflandırma sonuçları K-NN sınıflandırıcısının tüm indirgenmiş boyutlarda başarılı sonuçlar elde ettiğini göstermiştir. Ancak ÖH diğer boyut indirgeme yöntemlerine göre daha az başarı elde etmiştir. Buna sebep olarak giriş vektörünün boyutunun ikiye düşürülmesi ve bu boyutun sınıflandırma için tüm veriyi temsil etmede yetersiz kalması gösterilebilir.

Tablo 8. Giriş verisi olarak temel bileşenler için sınıflandırma sonuçları (%)

	Doğruluk	Duyarlılık	Özgünlük
<i>K-NN</i>	99.63	99.29	99.89
<i>YSA</i>	98.14	98.97	97.31
<i>DVM</i>	98.56	97.43	99.69
<i>KA</i>	95.90	93.23	98.57

Tablo 9. Giriş verisi olarak bağımsız bileşenler için sınıflandırma sonuçları (%)

	Doğruluk	Duyarlılık	Özgünlük
<i>K-NN</i>	99.26	98.80	99.71
<i>YSA</i>	97.79	99.06	96.51
<i>DVM</i>	98.04	97.00	99.09
<i>KA</i>	94.73	92.11	97.34

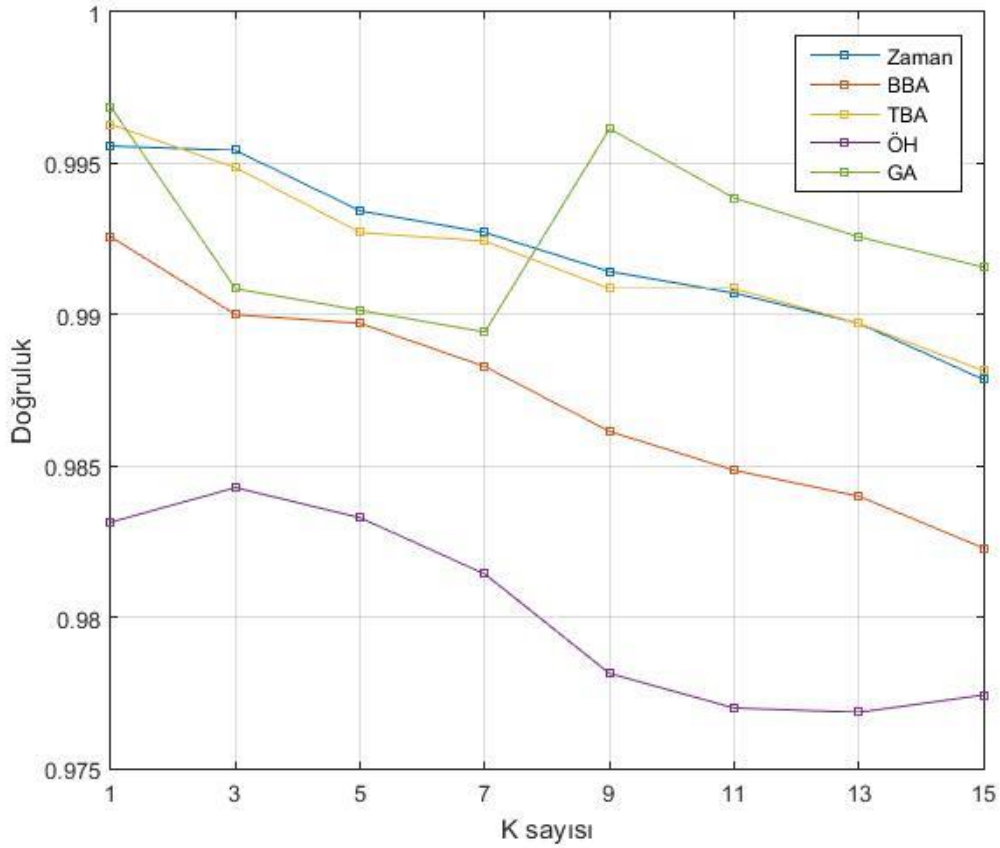
Tablo 10. Giriş verisi olarak ÖH tarafından üretilen öznitelikler için sınıflandırma sonuçları (%)

	Doğruluk	Duyarlılık	Özgünlük
<i>K-NN</i>	98.36	97.14	99.57
<i>YSA</i>	94.60	93.89	95.31
<i>DVM</i>	81.39	69.60	93.17
<i>KA</i>	77.54	55.29	99.80

Tablo 11. Giriş verisi olarak GA tarafından seçilen öznitelikler için sınıflandırma sonuçları (%)

	Doğruluk	Duyarlılık	Özgünlük
<i>K-NN</i>	99.69	99.46	99.91
<i>YSA</i>	98.53	98.91	98.14
<i>DVM</i>	98.10	96.89	99.31
<i>KA</i>	97.67	96.77	98.57

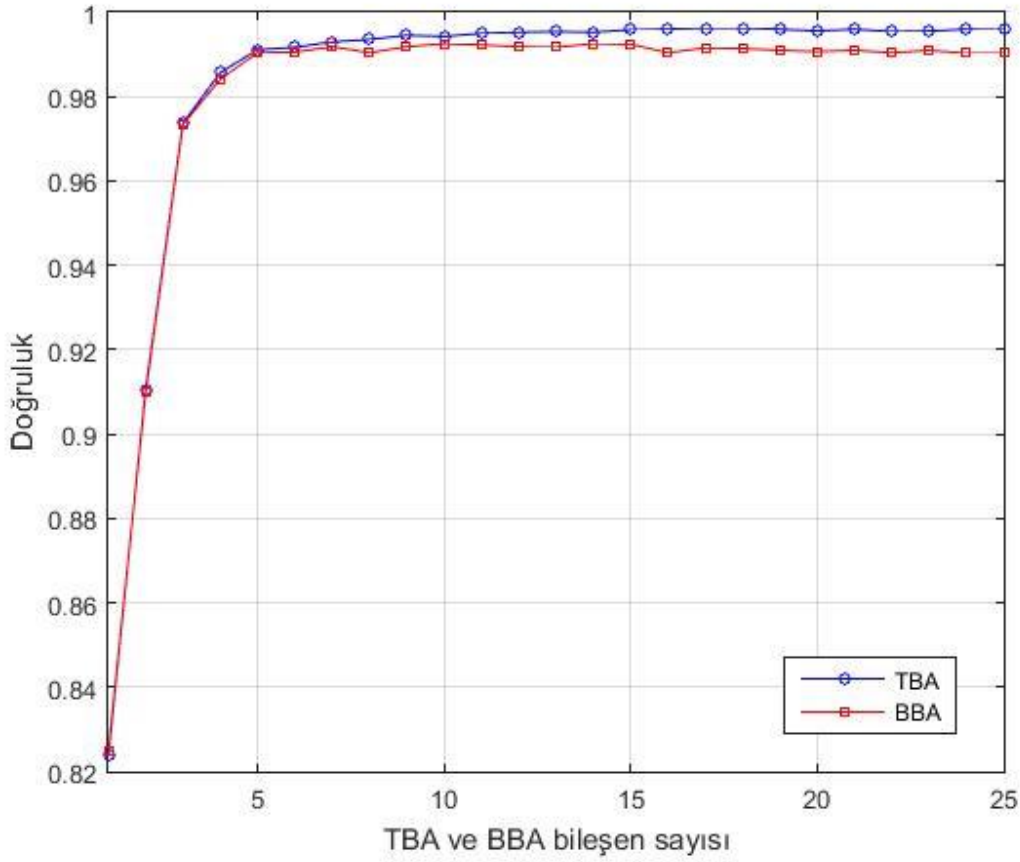
Testlerde görüldüğü gibi en iyi sınıflandırıcı olarak K-NN sınıflandırıcı, tüm giriş vektörlerinde öne çıkmaktadır. Bu nedenle bu sınıflandırıcı üzerinde en iyi sonucu verecek K değerini hesaplamak için parametre optimizasyonu yapılmıştır. En iyi K değerini bulmak için 1-15 aralığındaki tek sayılar sınıflandırıcıya K değeri olarak atanarak sınıflandırıcı başarımları test edilmiştir. En yüksek başarımları parametre olarak K=1 ve giriş vektörü olarak GA ile seçilmiş değerleri kullanan K-NN sınıflandırıcı **99,69% doğruluk** değeri ile elde etmiştir. K değerinin 1 olarak belirlenmesi örüntü tanıma işlemlerinde çok tercih edilmediğinden diğer K değerleri irdelenmiş ve sonuçların birbirine yakın değerler olduğu görülmüştür. Test aralığındaki tüm K değerleri, giriş vektörü olarak zaman serisi, TBA ve GA kullanılarak kabul edilebilir düzeyde yüksek başarımlar (>%98.8) elde etmiştir. Şekil 49'da K-NN sınıflandırıcısındaki K sayısının sınıflandırma başarımına etkisi gösterilmiştir. Özellikle K=9 için GA ile seçilmiş öznitelikler ile yapılan sınıflandırma sonucu kayda değer bir artış göstermiştir.



Şekil 49. K sayısının sınıflandırma başarımına etkisi

2.5.6. TB ve BB Sayısının Sınıflandırma Başarımına Etkisi

TBA ve BBA ile özniteliklerin hesaplanması aşamasında en iyi sonucu verecek bileşen sayılarını belirlemek için yapılan testlerde elde edilen veriler irdelenmiştir. Şekil 50'de Bağımsız Bileşen (BB) ve Temel Bileşen (TB) sayısının sınıflandırma başarımına etkisi gösterilmiştir. Sınıflandırıcı olarak K-NN sınıflandırıcısı kullanılmıştır.



Şekil 50. TB ve BB sayısının sınıflandırma başarımına etkisi

Şekilden de görülebileceği gibi TB ve BB sayısı beşe ulaştığında yeterli bir temsil gücü elde edilmiş oluyor ve 10'a ulaştığında ise artık başarımlar yatay olarak devam ediyor. Bu nedenle testlerde TB sayısı 17 olarak seçilmiş, BB sayısı ise 10 olarak belirlenmiş ve deneylerde bu şekilde kullanılmıştır.

2.5.7. Sınıflandırma Hesaplama Süreleri

Çalışmada kullanılan algoritmaların 10-kez çapraz doğrulama için test süreleri (sn.) ölçülmüş ve Tablo 12'de verilmiştir. Burada yine K-NN sınıflandırıcısı diğer sınıflandırıcılara göre üstünlük sağlamıştır. YSA mimarisi gereği özellikle giriş vektörünün boyutu arttığında yavaş kalmıştır. Boyut indirgeme algoritmalarının işlem hızına etkisi de açıkça görülmektedir.

Tablo 12. Sınıflandırıcıların kullandıkları giriş vektörüne göre sınıflandırma süreleri (sn)

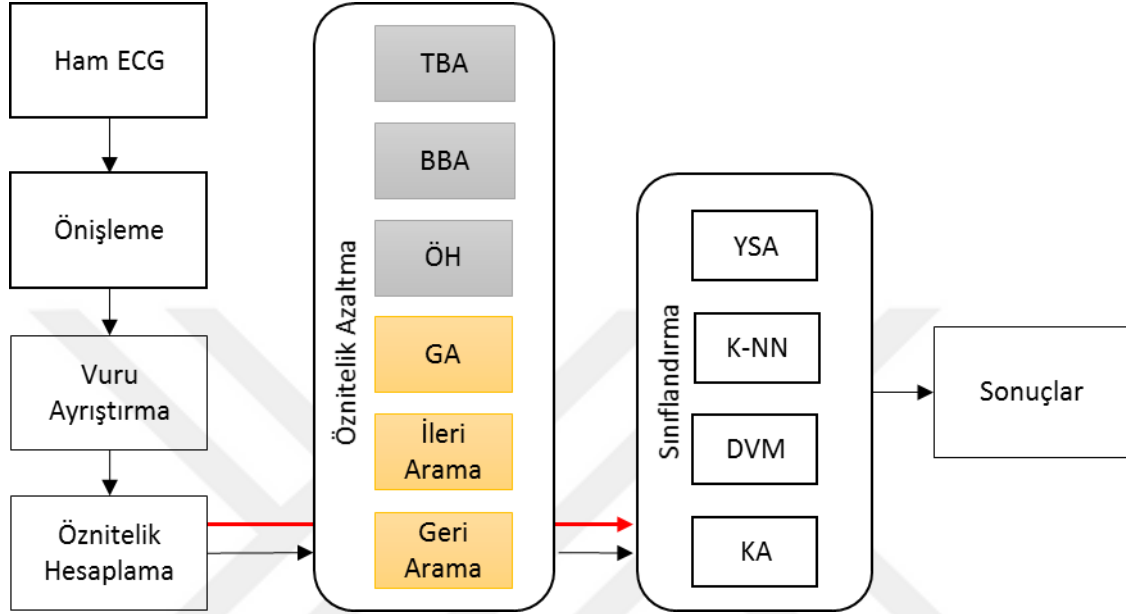
	Zaman	TBA	BBA	ÖH	GA
<i>K-NN</i>	45	4	3	1	10
<i>YSA</i>	2008	225	16	75	867
<i>DVM</i>	56	13	17	636	28
<i>KA</i>	181	21	19	7	17

2.6. EKG Vurularının Sınıflandırılması

Anormal kardiyak koşulların bilgisayar destekli uzun vadeli izleme sistemleri tarafından hızlı ve yeterli algılanması, farklı kalp vuru türleri arasında ayırım yapabilecek EKG vuru sınıflandırma sisteminin güvenilirliğine büyük ölçüde bağlıdır. Otomatik kalp vurusu sınıflandırma işlemi gerçekleştirecek olan yazılımın doğruluğu kardiyak işlev bozukluklarının tanısı, özellikle uzun süreli kayıtların (24 saat) izlenmesi veya Holter kayıtlarının analizinde büyük önem taşımaktadır. Kalp vuru analizi yöntemlerinin doğru performansı EKG işaretinin kalitesi, uygulanan sınıflandırma yöntemleri, EKG'yi tanımlayacak öznitelikler ve eğitimde kullanılacak veri kümesi dahil olmak üzere birçok önemli faktöre bağlıdır. EKG işaretinin kalitesinin iyileştirilmesi pratikte elektrotların güzel yerleştirilmesi, elektrot kablolarının hareketinin engellenmesi, işareti kaydeden ve gürültü gideren ekipmanın tasarımı vb. gibi parametrelere bağlıdır.

Bu çalışmada geliştirilen kalp vuru sınıflandırma şemasının blok diyagramı Şekil 51'de gösterilmiştir. Sistem incelendiğinde Şekil 45'te verilmiş EKK tanıma sistemiyle ön işleme ve vuru ayrıştırma aşamalarında benzerlikler içermektedir. Bu nedenle bir önceki bölümde bahsedilmiş olan ön işleme ve vuru ayrıştırma işlemlerinden bu bölümde tekrar bahsedilmeyecektir. Öznitelik hesaplama adımlarında bir vuruluk işarete ait istatistiksel öznitelikler hesaplanmaktadır. Çalışmada bir vuruluk işaretin genlik değerlerinden hesaplanmış olan öznitelikler için öznitelik azaltma işlemleri gerçekleştirilmiştir. Bu özniteliklerin kendileri ile birlikte boyutu azaltılmış sonuçları sınıflandırma algoritmalarında karşılaştırmalı olarak test edilmiş ve sonuçları irdelenmiştir. Öznitelik azaltma aşamasında TBA, BBA ve ÖH öznitelik çıkarımı yöntemleri ile öznitelik vektörünün boyutları azaltılmıştır. GA, SIA ve SGA algoritmaları ile özniteliklerden veriyi temsil yeteneği en iyi olanlar seçilmiştir. Bu aşamalarda en iyi sonucu elde etmek için parametre optimizasyon

işlemleri gerçekleştirilmiştir. Sınıflandırma adımında YSA, K-NN, DVM ve KA algoritmaları deneylerde karşılaştırılmış ve en iyi sonucu verecek sınıflandırma şeması belirlenmiştir.



Şekil 51. EKG vuru sınıflandırma deneyinin genel blok diyagramı

EKG arařtırmacıları için son yıllarda standart haline geldiđi için deneylerde kullanılan tüm kayıtlar MIT-BIH aritmi veritabanından alınmıřtır. Hesaplamalarda MATLAB paket programı kullanılmıřtır. Geliřtirilen vuru sınıflandırma řemalarının performansları birbirileri ile ve literatürdeki diđer çalıřmalar ile *dođruluk*, *özgünlük*, *seçicilik* ölçütleri kullanılarak karşılaştırılmıřtır.

2.6.1. Testlerde Kullanılan Veri Seti

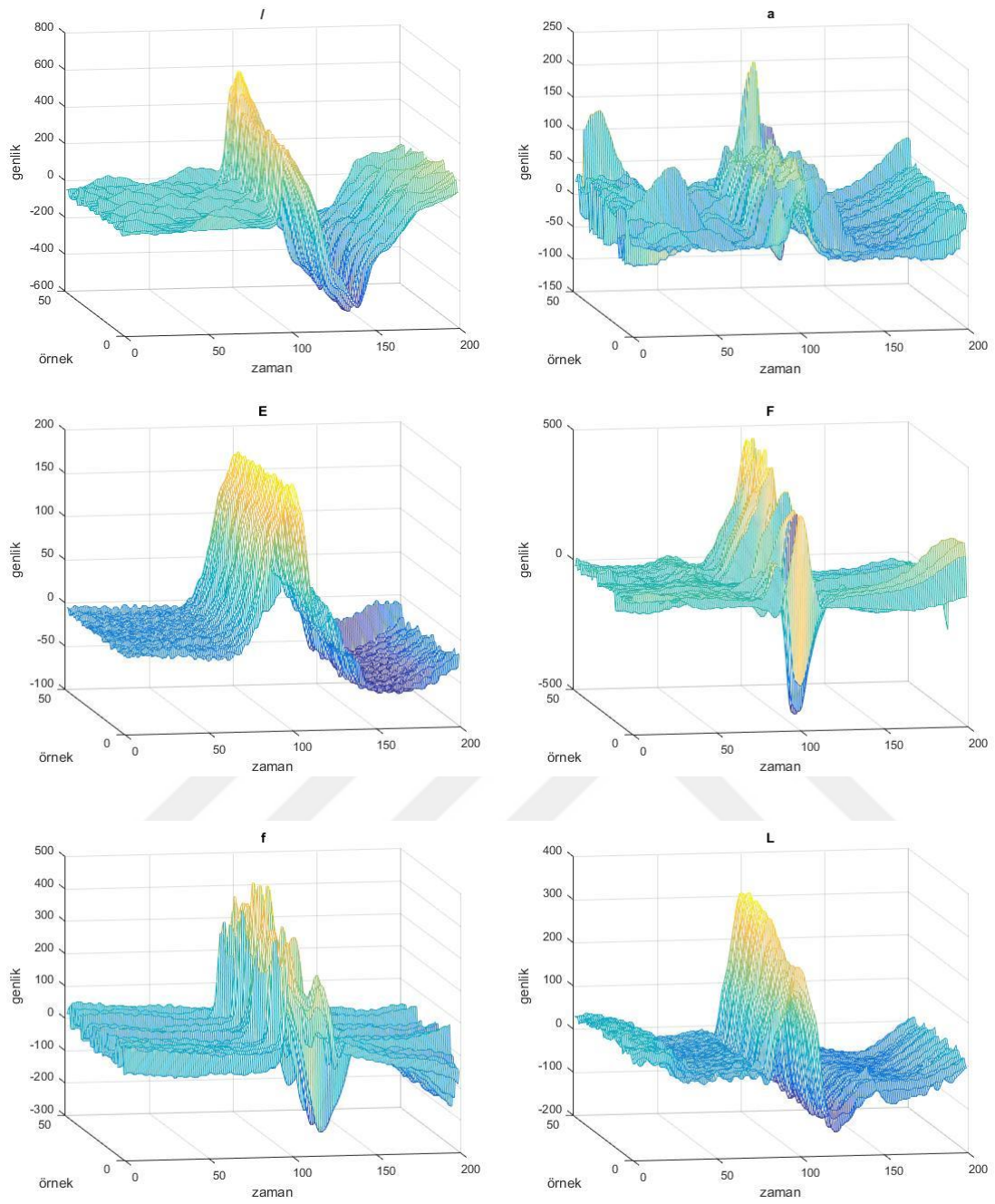
Testlerde MIT-BIH aritmi veritabanında bulunan en yaygın dokuz vuru türünde veriler seçilmiř ve bu dokuz sınıfta sınıflandırma işlemleri gerçekleştirilmiřtir. Bu vurular Normal vuru (N), sol dal blok vurusu (L), sađ dal blok vurusu (R), erken karıncık kasılması (V), yapay vuru (P veya /), karıncık ve normal vuru birleřimi (F), yapay ve normal vuru birleřimi (f), anormal kulakçık erken vurusu (a), karıncık kaçak vurusudur (E). Tablo 13'te vuru türleri ve bu vuru türlerine ait sayılar gösterilmiřtir. Çalıřmada kullanılacak vuru sayıları

belirlenirken sayısı 1000'den küçük olan vuruların tamamı testlere dahil edilmişken, sayısı 1000'den büyük olan vurular için 1000 adet vuru seçilmiştir.

Tablo 13. Testlerde kullanılan vuru türleri ve örnek sayıları

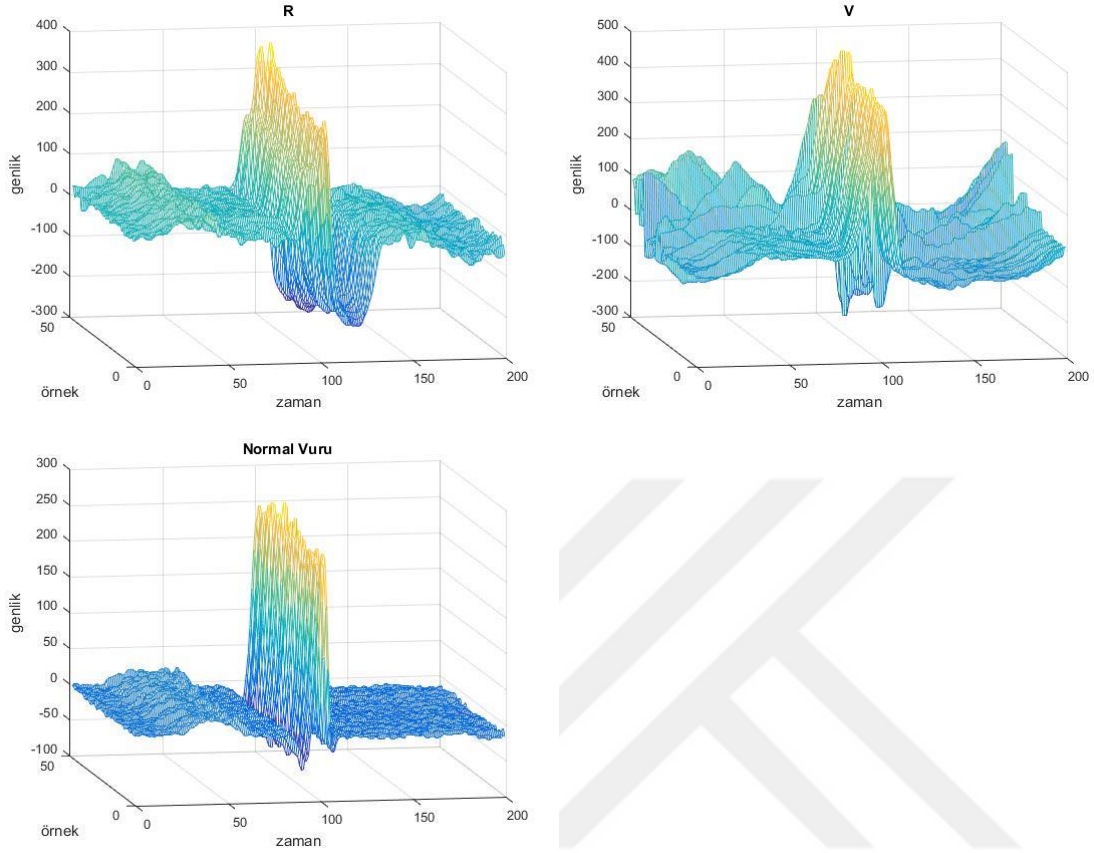
Sınıf	Vuru Sayısı
N	1000
L	1000
R	1000
V	1000
/	1000
F	802
f	260
a	150
E	106
Genel Toplam	6318

Şekil 52'de veri setindeki her bir vuru tipine ait ilk 50 örnek için grafiksel gösterim sunulmuştur.



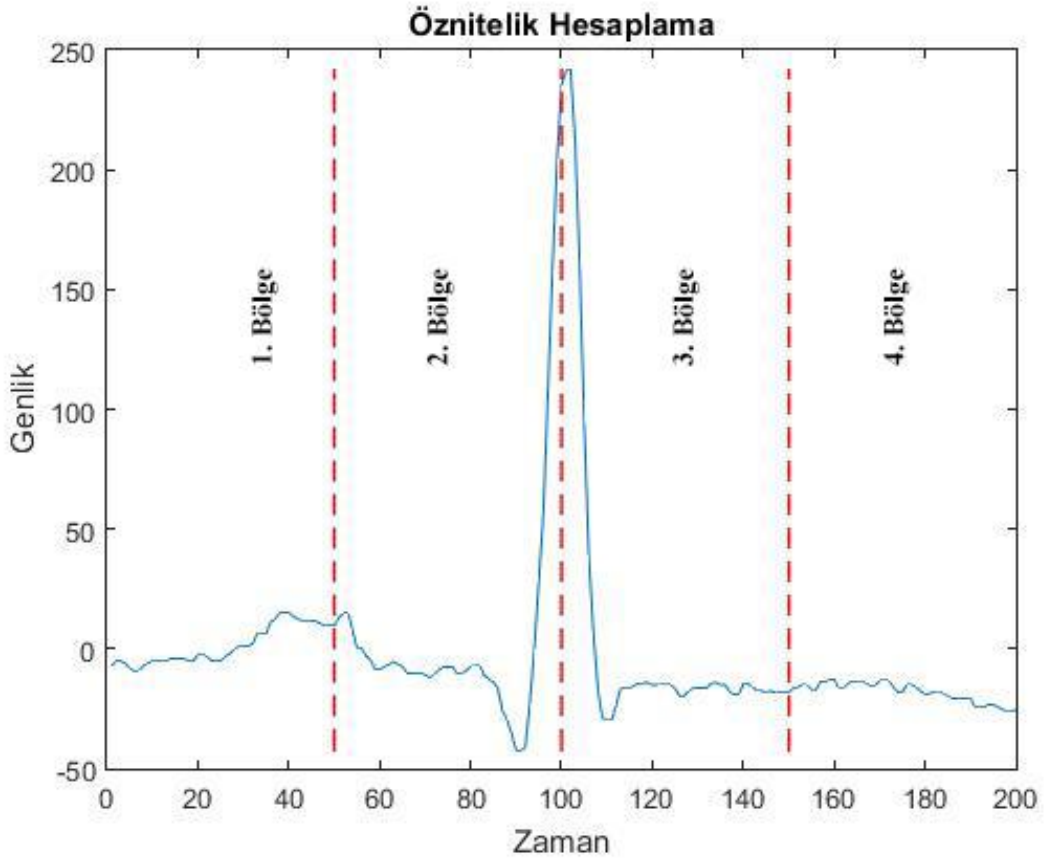
Şekil 52. Veri setindeki her bir vuru tipine ait ilk 50 örnek

Şekil 52'nin Devamı



2.6.2. Öznitelik Hesaplama

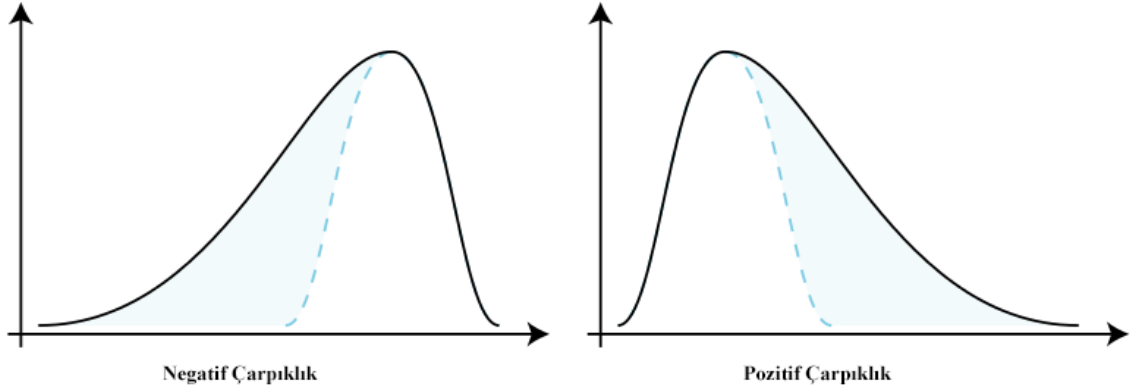
Vuru ayrıştırma aşamasından sonra sınıflandırma işlemlerinde vuruyu temsil edecek öznitelikler hesaplanmaktadır. Bir vuruya ait 200 örnek noktadan oluşan genlik değerleri vuru ayrıştırma aşamasında elde edilmiştir. Bu değerlere ait aşağıda açıklanan çarpıklık, basıklık, dağılım aralığı, çeyrekler aralığı, standart sapma ve ortalama istatistiksel öznitelikleri sınıflandırma şemalarında kullanılmak üzere hesaplanmıştır. Bu değerler ilk olarak bir vuruluk 200 noktadan oluşan veri üzerinde hesaplanmıştır. Özniteliklerin veriyi temsil yeteneğini arttırmak için Şekil 53'de gösterildiği gibi vuru eşit genişlikli 4 ayrı bölgeye ayrılmış ve her bölge için aynı hesaplamalar gerçekleştirilmiştir. Bu özniteliklere ek olarak hesaplanan vuruya ait önceki ve sonraki R-R aralıkları (örnek sayısı) da öznitelik olarak hesaplanmıştır.



Şekil 53. Vuruya ait özniteliklerin hesaplanması

Çarpıklık ve Basıklık

Olasılık teorisi ve istatistikte çarpıklık, gerçek değerli rasgele değişkenin olasılık dağılımının ortalama civarındaki simetrisini gösteren bir ölçüttür. Çarpıklık değeri pozitif, negatif veya sıfır değerini alabilir. Çarpıklık değerinin sıfır olması dağılımın simetrik olduğunun bir göstergesidir. Şekil 54’de pozitif ve negatif çarpıklığa ait grafikler gösterilmiştir.



Şekil 54. Negatif ve Pozitif Çarpıklık

Çarpıklık üçüncü standartlaştırılmış moment olarak ifade edilir ve matematiksel gösterimle eşitlik (68)'deki gibi tanımlanır.

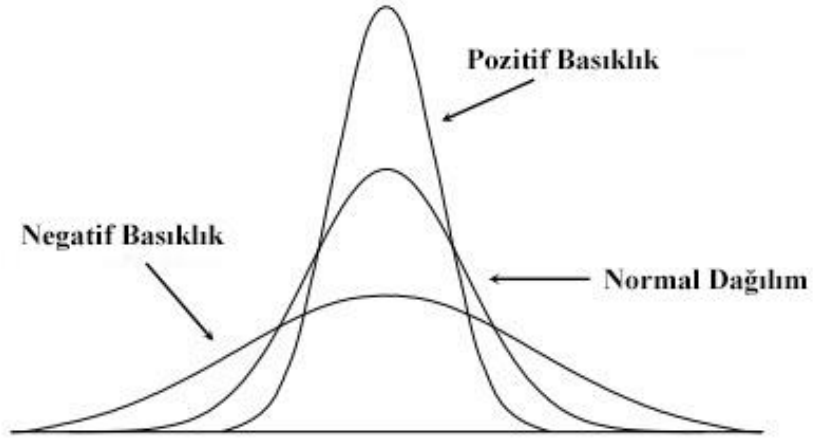
$$C = \frac{\mu_3}{\sigma^3} \quad (68)$$

Burada μ_3 üçüncü ortalama civarındaki moment olarak ve σ standart sapma olarak ifade edilir. n sayıda örnek içeren bir örneklem için çarpıklık eşitlik (69)'daki gibi hesaplanır.

$$C = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(\sum_{i=1}^n (x_i - \bar{x}))^3} \quad (69)$$

Burada x_i i . örnek değerini ve \bar{x} ise örneklem aritmetik ortalamasını göstermektedir.

Basıklık kavramı bir gerçek değerli rasgele değişken için olasılık dağılımının sivriliğini veya basıklığını gösteren bir ölçümdür. Bir diğer deyişle bir dağılımın normal dağılış eğrisine göre tepeli veya basık olma durumudur. Basıklığın pozitif olması sivriliğin fazla olduğu, negatif olması ise sivriliğin daha düz olduğu anlamına gelir. Şekil 55'de basıklık kavramı gösterilmiştir.



Şekil 55. Basıklık gösterimi

Basıklık dördüncü standartlaştırılmış moment olarak ifade edilir ve matematiksel gösterimle eşitlik (70)'deki gibi tanımlanır.

$$B = \frac{\mu_4}{\sigma^4} \quad (70)$$

Burada μ_4 dördüncü ortalama civarındaki moment olarak ve σ standart sapma olarak ifade edilir. n sayıda örnek içeren bir örneklem için basıklık eşitlik (71)'deki gibi hesaplanır.

$$B = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(\sum_{i=1}^n (x_i - \bar{x}))^4} \quad (71)$$

Burada x_i i . örnek değerini ve \bar{x} ise örneklem aritmetik ortalamasını göstermektedir. Çalışmada öznitelik olarak her vuru ve vuruya ait 4 bölüm için ayrı ayrı çarpıklık ve basıklık değerleri hesaplanmıştır.

Diğer İstatistiksel Özniteliklerin Hesaplanması

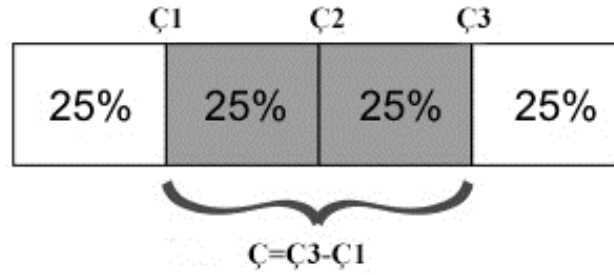
Her bir vuru ve vuruya ait 4 bölüm için dağılım aralığı, çeyrekler aralığı, standart sapma ve ortalama istatistiksel öznitelikleri hesaplanmıştır. Bu istatistikler aşağıdaki (72)-(75) eşitliklerine göre hesaplanırlar.

$$A = \text{Maksimum}(x_i) - \text{Minimum}(x_i) \quad (72)$$

Burada x_i i . örnek değerini göstermektedir ve dağılım aralığı bu değerlerin aldığı en büyük ve en küçük değerlerin farkı olarak hesaplanmaktadır. Çeyrekler aralığı, istatistiksel olarak dağınıklığın bir ölçütüdür ve alt ve üst çeyreklikler arasındaki farka eşittir.

$$\zeta = \zeta_3 - \zeta_1 \quad (73)$$

Burada ζ_1 , ilk çeyrekliği gösterir ve sıralanmış veride ilk %50'de kalan alandaki ortancaya eşittir. ζ_3 ise üçüncü çeyreklik anlamına gelir ve sıralanmış veride ikinci %50'de kalan verinin ortanca değerine eşittir. Şekil 56 çeyrekler aralığı için sıralanmış verilerde hesaplanmasını göstermektedir.



Şekil 56. Çeyrekler aralığının hesaplanması

Standart sapma veri değerlerinin istatistiksel olarak dağılımını gösteren bir ölçüdür ve varyansın kareköküne eşittir.

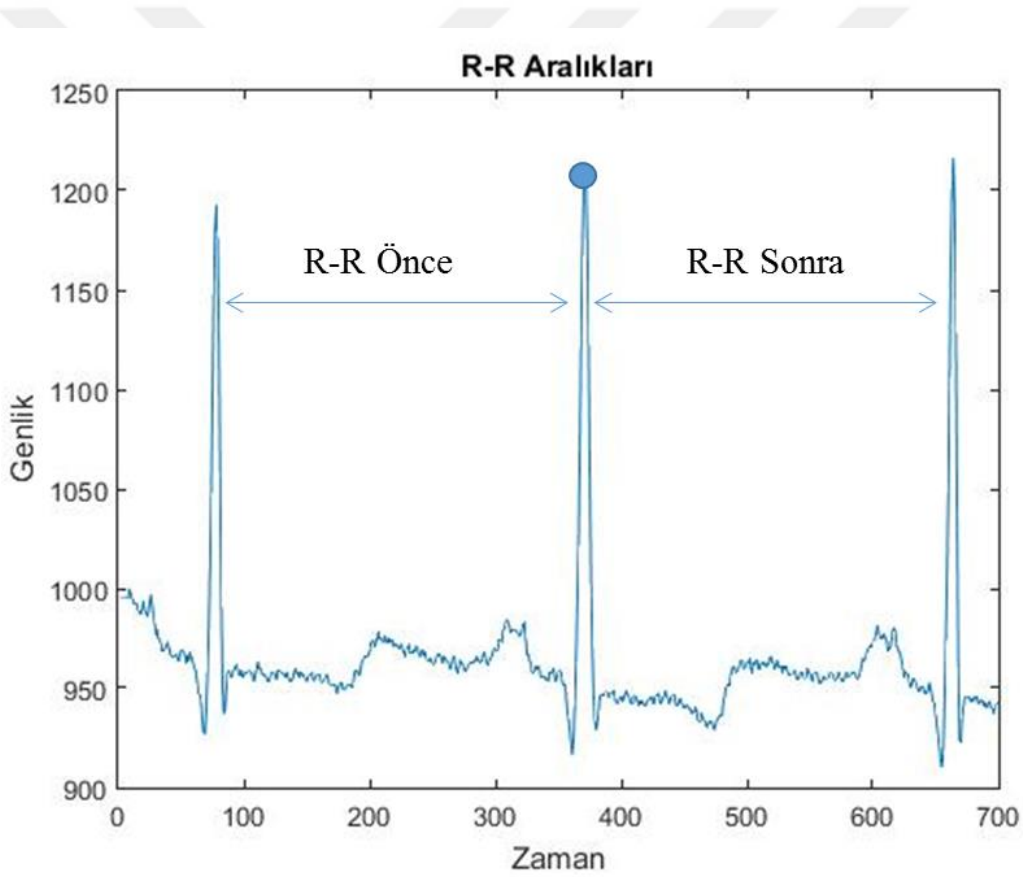
$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (74)$$

Burada x_i i . örnek değerini, \bar{x} örneklem aritmetik ortalamasını ve n örnek sayısını göstermektedir. Ortalama veri değerlerini için tek bir orta değerini temsil eden istatistiksel bir ölçüdür.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (75)$$

RR Aralıklarının Hesaplanması

Hesaplamalarda aktif vurunun bir önceki ve sonraki vurular ile olan süreleri de hesaplanarak öznitelik olarak kullanılmıştır. Bu amaçla R-R Önce ve R-R Sonra değerleri hesaplanmıştır. Bu değerler Şekil 57’de görüldüğü gibi bir önceki ve bir sonraki vuruya ait R tepesine olan örnek sayıları olarak hesaplanmıştır.

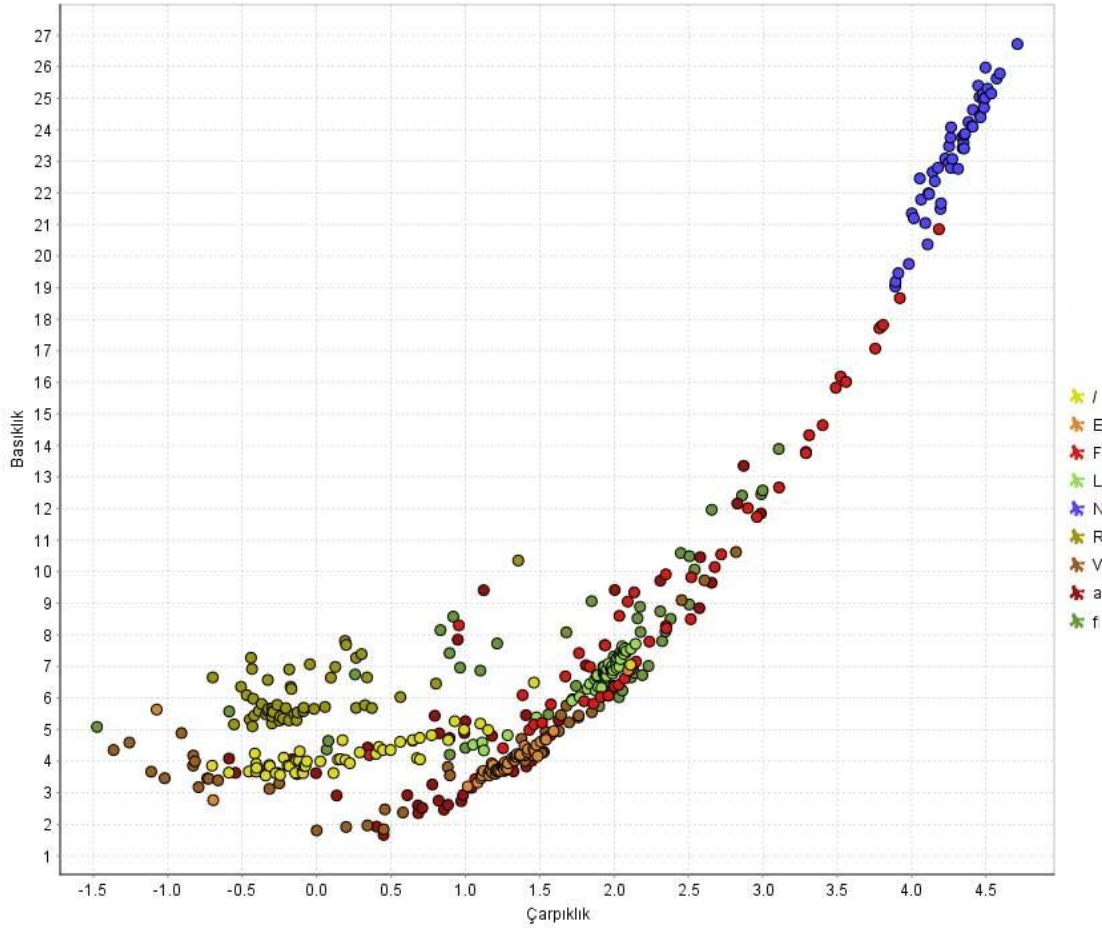


Şekil 57. R-R Aralıklarının hesaplanması

Böylece öznitelik olarak her vuru ve 4 alt bölümü için 6 istatistik hesaplanmıştır. Bunlara ek olarak iki R-R aralığı bu özniteliklere eklenerek öznitelik vektörünün boyutu 32 olarak belirlenmiştir.

Tablo 14. Hesaplanan Öznitelikler

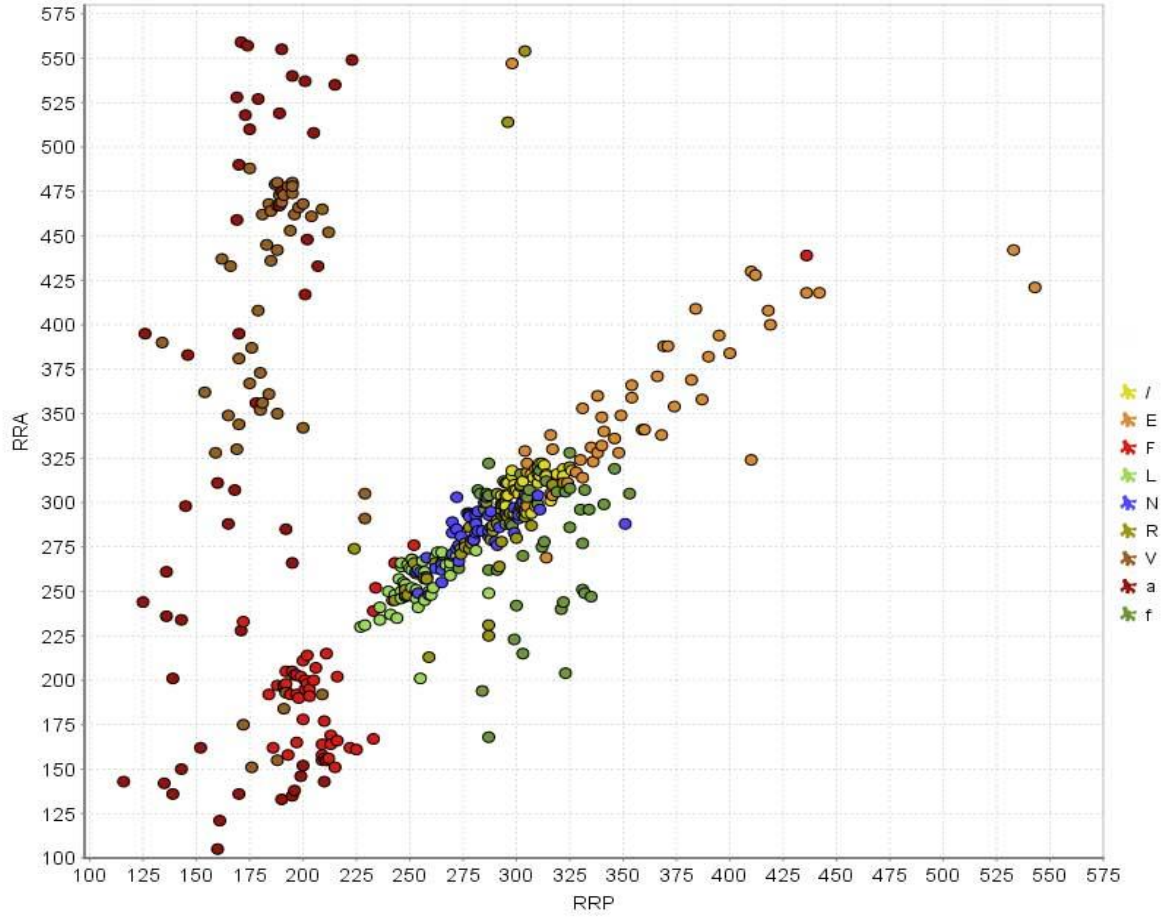
İstatistiksel Öznitelikler						
Öznitelik	Çarpıklık	Basıklık	Aralık	Çeyreklik	St. Sapma	Ort.
Tüm Vuru	SKEW0	KURT0	RANG0	IQR0	STD0	MEA0
Bölge 1	SKEW1	KURT1	RANG1	IQR1	STD1	MEA1
Bölge 2	SKEW2	KURT2	RANG2	IQR2	STD2	MEA2
Bölge 3	SKEW3	KURT3	RANG3	IQR3	STD3	MEA3
Bölge 4	SKEW4	KURT4	RANG4	IQR4	STD4	MEA4
R-R Aralığı						
Önceki R-R	RRP					
Sonraki R-R	RRA					



Şekil 58. Bir vuruluk işaret için Çarpıklık ve Basıklık öznitelikleri

Bu öznitelik vektörü sınıflandırma şemalarında testlerde karşılaştırılmalı olarak kullanılmıştır. Şekil 58’de bir vuruluk işaretin tamamı kullanılarak hesaplanan *çarpıklık* ve *basıklık* öznitelikleri gösterilmiştir. Grafiğin daha anlaşılır olması için her sınıftan 50 örnek

seçilmiştir. Benzer şekilde aktif vurudan önceki R-R aralığı ve sonraki R-R aralığı özniteliklerinin dağılımları Şekil 59’da gösterilmiştir.



Şekil 59. Önceki R-R aralığı ve Sonraki R-R Aralığı öznitelikleri

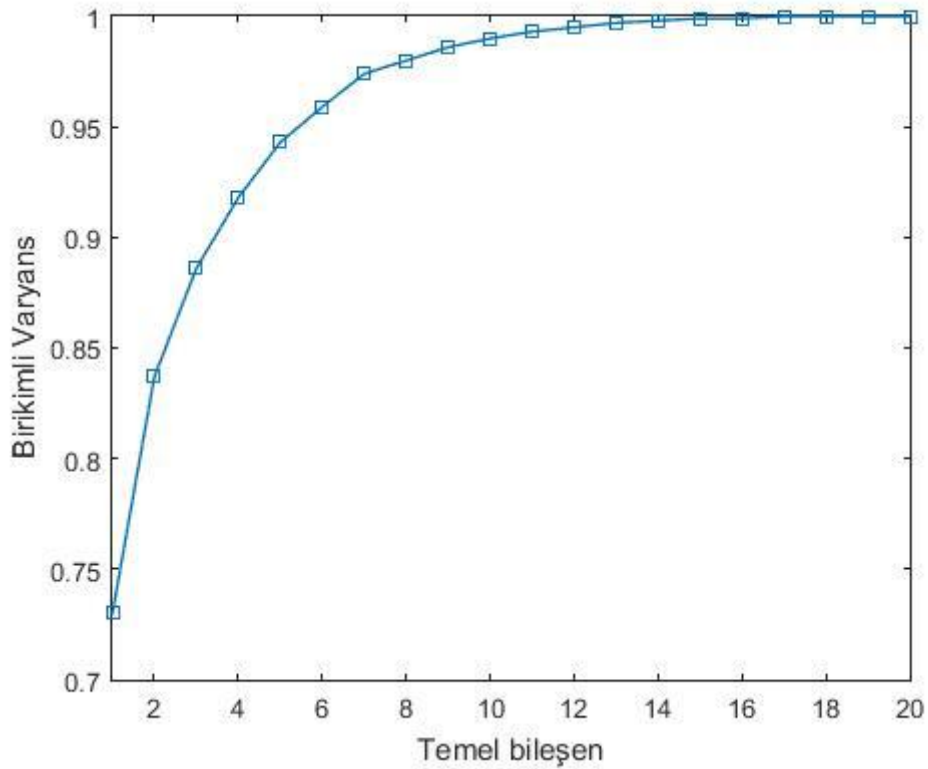
2.6.3. EKG Vurularının Sınıflandırılması İçin Öznitelik Boyut İndirgeme

Bir önceki konuda 32 adet öznitelik hesaplanarak sınıflandırma için giriş vektörü oluşturulmuştu. Sınıflandırma aşamasına geçilmeden önce giriş verisinin daha iyi temsil edilmesi için öznitelik boyut indirgeme işlemleri gerçekleştirilmiştir. Burada boyut indirgeme işlemi iki farklı teknik ile gerçekleştirilecektir. İlk yöntemde TBA, BBA ve ÖH kullanılarak bulunan öznitelik vektörleri farklı uzaya taşınarak verinin temsili sağlanmıştır. Bu yöntemde bir vuruluk işaretinden elde edilmiş özniteliklerden yeni öznitelikler farklı hesaplama teknikleri kullanılarak hesaplanacak ve boyut indirgenecektir.

İkinci yöntem ise elde bulunan özneliklerden veriyi en iyi temsil edecek olanların seçilmesidir. Bu kısımda SİA, SGA ve GA yöntemi özneliklerin seçimi için kullanılacaktır.

TBA, BBA ve ÖH Kullanılarak Boyut İndirgeme

Bu aşamada ilk olarak veriyi en iyi temsil edecek TB sayısını belirlemek için ızgara arama yöntemi kullanılarak en uygun TB sayısı belirlenmiştir. Arama işlemlerinde özneliklerin değerlendirilmesi için bir sınıflandırıcı değerlendirme fonksiyonu olarak kullanılmış ve her bir adımda sınıflandırma doğruluğu kontrol edilerek TB sayısı ve karşılık gelen sınıflandırma başarımı araştırılmıştır. Bu amaçla 1-32 arasındaki TB'ler hesaplanarak her adımda sınıflandırma sonuçları değerlendirilmiştir.



Şekil 60. TB'in birikimli varyans değerleri

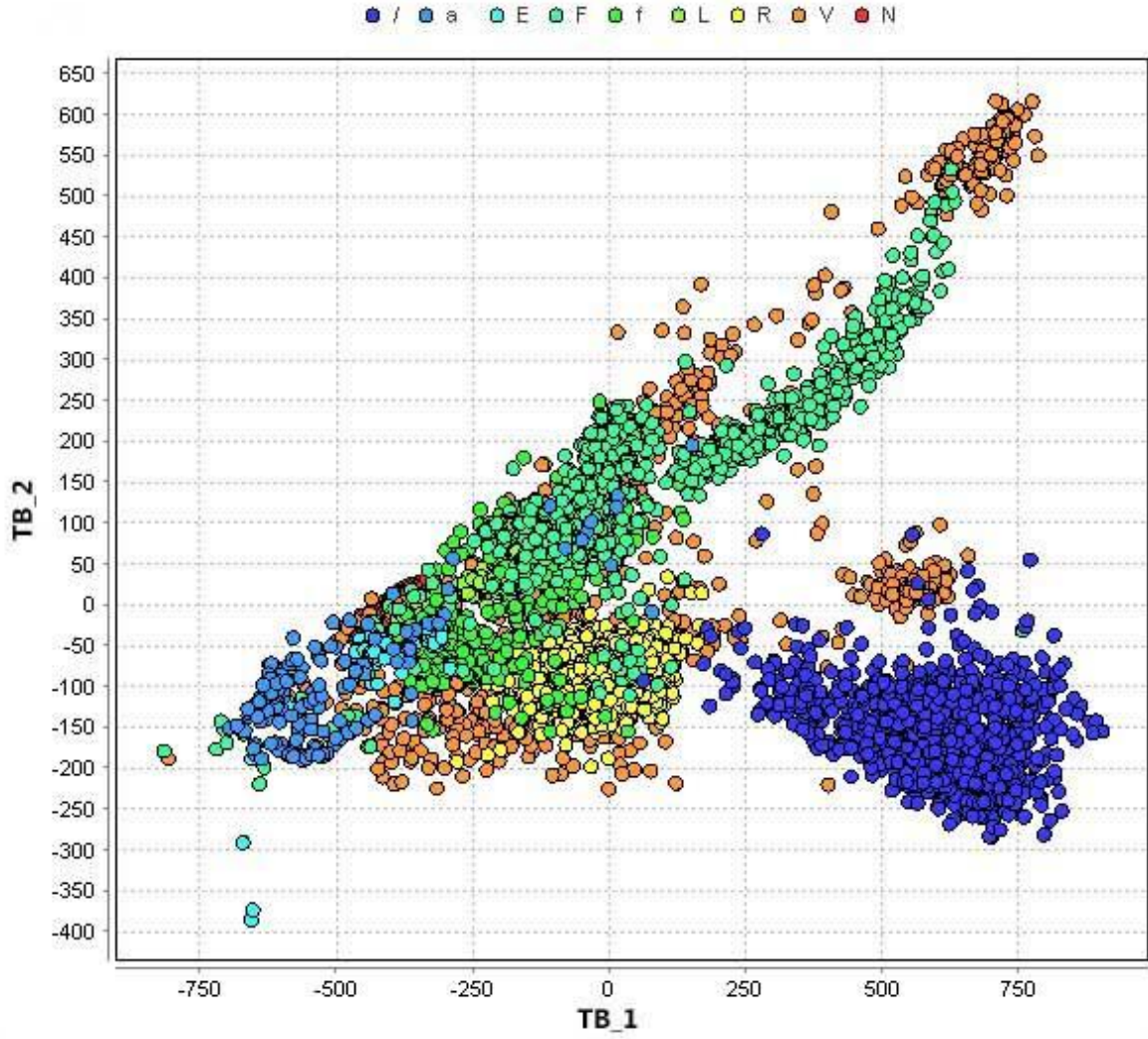
Şekil 60'ta hesaplanan ilk 20 temel bileşene ait birikimli varyans değerleri gösterilmiştir. Hesaplanan ilk iki TB'in veriyi temsil edecek en önemli bileşenler olduğu TBA konusu anlatılırken açıklanmıştı. İlk iki temel bileşene ait birikimli varyans değeri %83,7'dir. Yani ilk iki bileşen tüm veriyi temsil etme kabiliyeti olarak yaklaşık %84'lik bir

değeri içerisinde barındırmaktadır. Birikimli varyans değerleri TB sayısı 10 olduğunda ise %99 değerine ulaşmaktadır.

Tablo 15. TB sayısı ve sınıflandırma başarımı

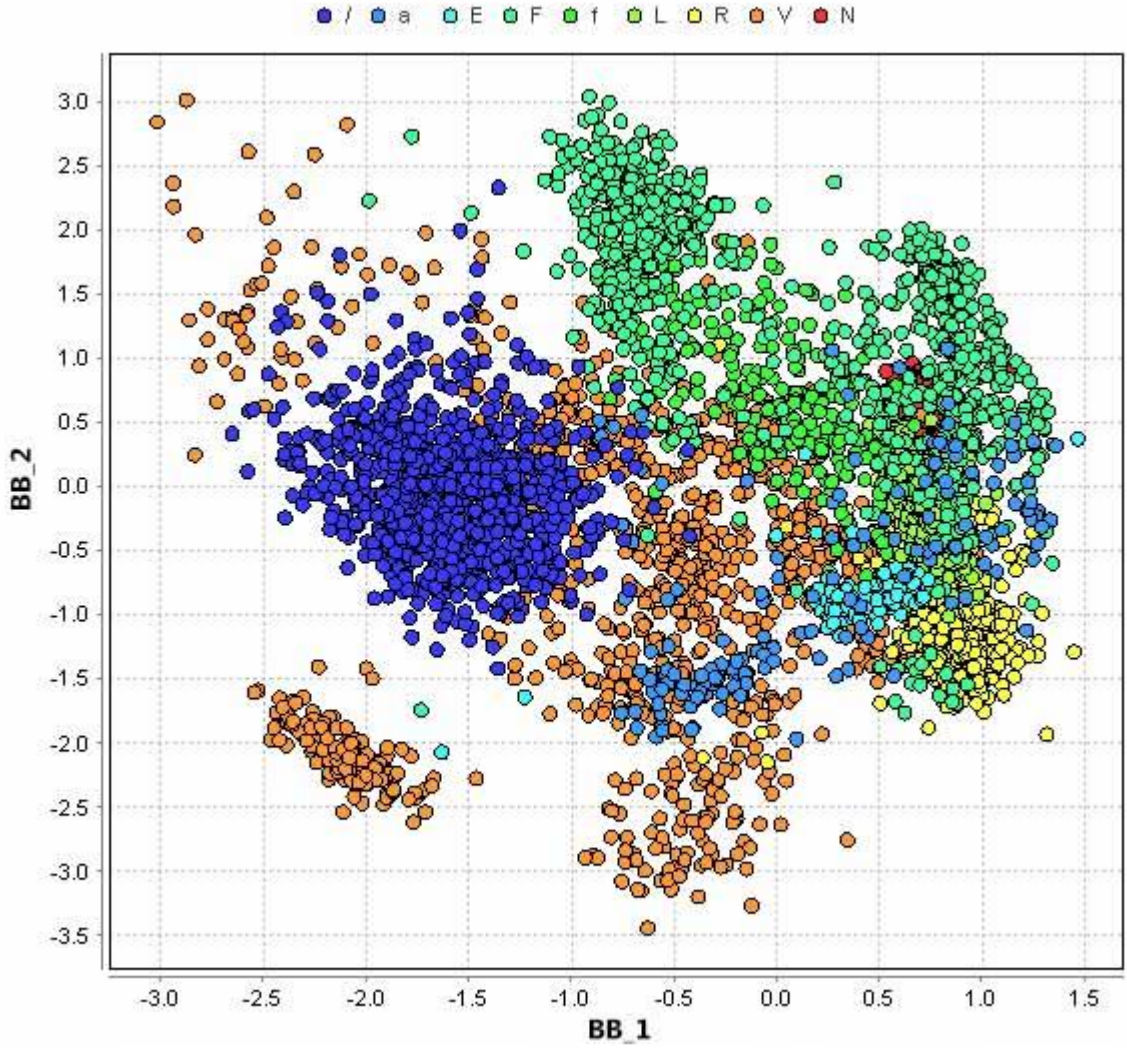
TB Sayısı	Sınıflandırma Başarımı (%)
1	46,66
2	77,98
3	88,86
4	95,65
5	97,48
6	98,35
7	98,51
8	98,84
9	98,83
10	98,89
11	98,91
12	98,86
13	98,91
14	98,97
15	98,97
16	98,92
17	99,00
18	98,96
19	98,96
20	98,96
21	99,00
22	98,97
23	98,99
24	98,96
25	98,94

Tablo 15'te ilk 25 temel bileşen için sınıflandırma başarımları verilmiştir. Görüldüğü gibi TB sayısı 17 olduğunda en yüksek sınıflandırma başarımı olan %99 değerine ulaşılmıştır. Bu nedenle sonraki testlerde TB sayısı 17 olarak alınacaktır. Şekil 61'de ilk iki temel bileşen için öznitelik değerleri gösterilmiştir.



Şekil 61. İlk iki temel bileşenin test verisinin temsili

Çalışmada diğer sonuçlar ile karşılaştırmak için hesaplanan özneliklerin BB'leri hesaplanmıştır. 1-32 arasındaki tüm BB'ler, en iyi sonucu verecek BB sayısını bulmak için teker teker hesaplanmış ve sınıflandırma başarımları değerlendirilmiştir. Testler sonucunda BB sayısı 9 olarak hesaplanmış ve bundan sonraki hesaplamalarda BB sayısı 9 olarak kullanılmıştır. Şekil 62'de ilk iki bağımsız bileşen için öznelikler gösterilmiştir.



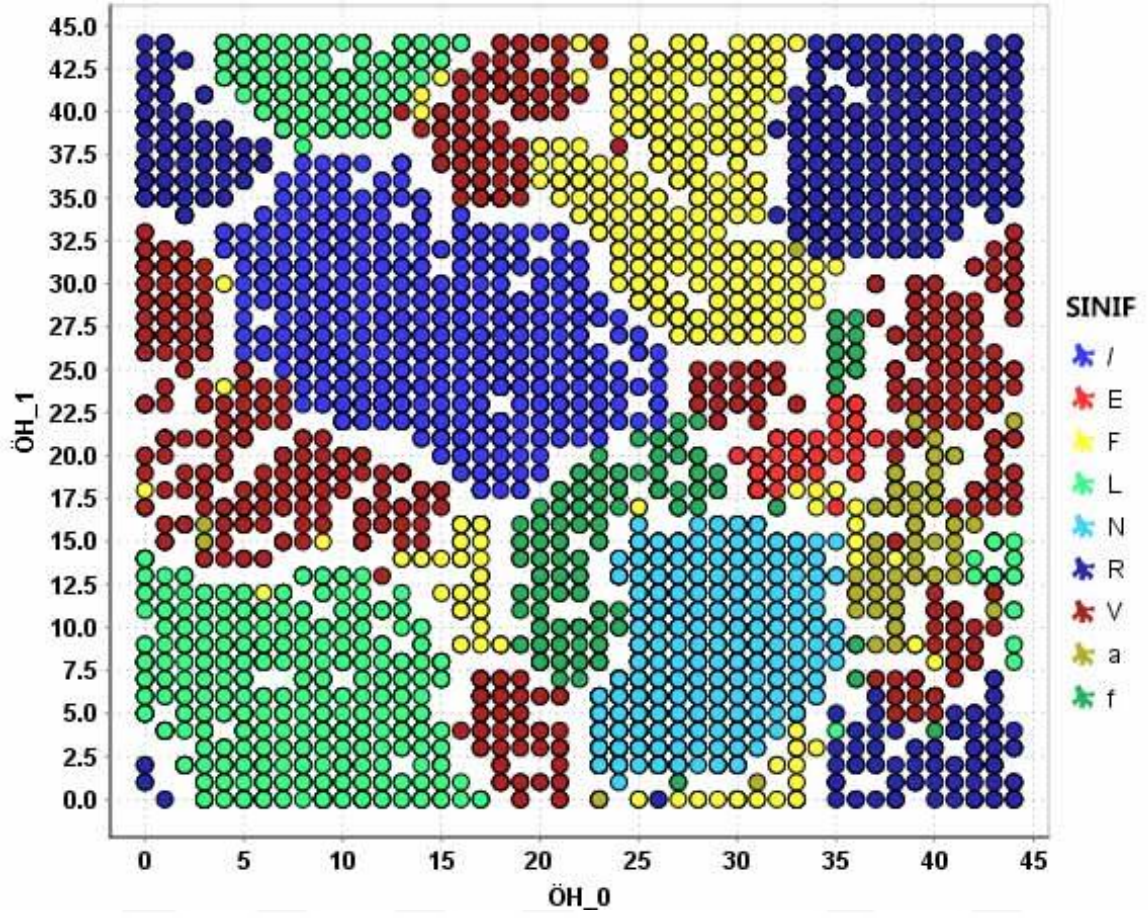
Şekil 62. İlk iki bağımsız bileşenin test verisini temsili

Tablo 16’da BB sayısına karşılık gelen sınıflandırma başarımları gösterilmiştir. BB’ler hesaplanırken hızlı BBA algoritması kullanılmıştır. Hesaplama neg-entropi yaklaşım fonksiyonu olarak logcosh, alfa parametresi olarak 1, maksimum iterasyon sayısı olarak 200 değerleri kullanılmıştır. Tablo incelendiğinde BB’lerin sınıflandırma başarımları test edilirken yaklaşık 10 civarında en yüksek sınıflandırma başarımlarını elde ettiği ve BB sayısı 10’u geçtiğinde yaklaşık aynı sınıflandırma başarımlarının elde edildiği görülmüştür. Bu nedenle testlerde bağımsız bileşen sayısının 10’dan büyük bir değer seçilmesinin sisteme herhangi bir pozitif katkısı olmayacaktır.

Tablo 16. BB Sayısı ve sınıflandırma başarımı

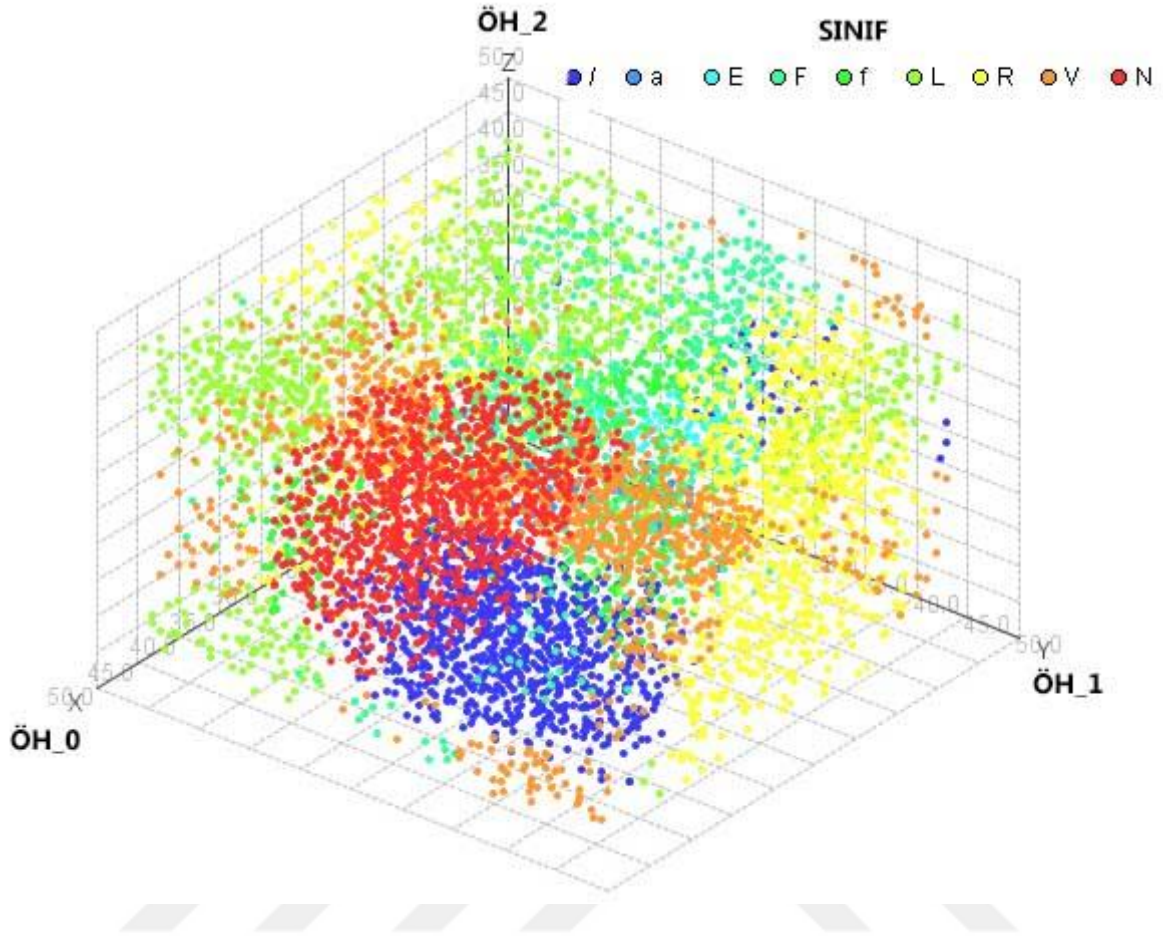
BB Sayısı	Sınıflandırma Başarımı (%)
1	46,95
2	78,68
3	88,89
4	95,46
5	97,53
6	98,61
7	98,73
8	98,97
9	99,21
10	99,11
11	99,11
12	99,13
13	99,13
14	99,07
15	99,08
16	99,05
17	99,05
18	99,18
19	99,02
20	99,11
21	99,05
22	99,16
23	99,10
24	99,03
25	99,00

Çalışmada kullanılan bir diğer boyut indirgeme yöntemi ÖH'dır. ÖH'ın en iyi sonuç verdiği değerleri hesaplamak için ÖH ağının çıkışı ve ağın boyutu farklı değerlerde test edilerek parametre optimizasyonu yapılmıştır. Bu optimizasyon aşamasında ağın çıkış boyut sayısı 2 ve 3, ağın genişliği ise 30-60 değerleri arasındaki değerlerde testlere tabi tutulmuştur. Ağın çıkış sayısı 3 ve ağ boyutu 45 parametreleri için en iyi sonuçları elde etmiştir. Ağın eğitimi için öğrenme tur sayısı olarak 30, başlangıçtaki öğrenme katsayısı 0,8, bitiş öğrenme katsayısı 0,01, başlangıçtaki (ilk turdaki) daire çapı 10, ve son tur daire çapı 1 olarak belirlenmiştir. Veri setinin ÖH ile iki boyutlu uzaya dönüştürülmüş şekli Şekil 63'de gösterilmiştir.



Şekil 63. ÖH kullanılarak öznitelik uzayının iki boyuta indirgenmiş hali

Benzer şekilde öznitelik uzayının üç boyutlu uzaya indirgenmiş hali Şekil 64'de gösterilmiştir.



Şekil 64. ÖH kullanılarak öznelik uzayının 3 boyutlu uzaya indirgenmiş hali

GA, SİA ve SGA Yöntemleri ile Öznelik Seçme

Önceki başlıkta istatistiksel yaklaşımlar uygulanarak yeni öznelikler hesaplanmış ve bunların içerisinde veriyi daha az boyutta temsil edebilecek yeni öznelik uzayları tanımlanmıştı. Bu başlık altında ise orijinal öznelikler arasından veriyi en iyi temsil edeceklerin seçimi gerçekleştirilmektedir.

Bu amaçla ilk olarak GA 32 öznelikten oluşan giriş verisinin boyutunu azaltmak için kullanılmıştır. Bu çalışmada turnuva temelli seçim stratejisi kullanan standart bir GA kullanılmıştır. Testlerde kullanılan GA parametreleri deneysel sonuçlar ile elde edilmiş ve kullanılan parametre değerleri aşağıda belirtilmiştir;

- Popülasyon boyutu: 5
- Maksimum jenerasyon sayısı: 50

- Çaprazlama olasılığı: 0,5
- Çaprazlama tipi: Uniform
- Bir özelliğin ilk durumdaki seçilme olasılığı: 0,5
- Mutasyon Olasılığı: 1/32 (32 giriş öznitelik sayısı)
- En az seçilebilecek öznitelik sayısı: 5

Burada giriş öznitelik sayısını hesaplanabilir boyutta (32) olduğu için herhangi bir üst sınır eklenmemiştir. Ancak işlem yükünün bir miktar azaltılabilmesi için alt sınır olarak 5 değeri seçilmiştir. GA ile seçim işlemi gerçekleştirildikten sonra 17 adet öznitelik seçilmiştir. Bu öznitelikler Tablo 17’de verilmiştir. Tabloda Seçim Durumu 1 olan öznitelikler GA tarafından seçilmiş özniteliklerdir [27].

Tablo 17. GA tarafından seçilen öznitelikler

Öznitelik	Seçim Durumu	Öznitelik	Seçim Durumu
SKEW0	0	STD2	1
KURT0	0	MEA2	1
RANG0	0	SKEW3	1
IQR0	1	KURT3	0
STD0	0	RANG3	0
MEA0	1	IQR3	1
SKEW1	0	STD3	1
KURT1	1	MEA3	1
RANG1	1	SKEW4	0
IQR1	0	KURT4	0
STD1	1	RANG4	0
MEA1	1	IQR4	1
SKEW2	1	STD4	0
KURT2	0	MEA4	1
RANG2	0	RRP	1
IQR2	0	RRA	1

SİA ve SGA yöntemleri kaynak tüketici yöntemler olarak bilinir ve giriş öznitelik boyutu büyüdüğünde bu yöntemleri uygulamak imkânsız hale gelir. İlk bölümde erken karıncık kasılması vurusunun algılanmasında giriş vektörünün boyutu 200 olduğu için bu yöntemler uygulanamamıştır. Burada giriş öznitelik boyutu 32 olduğu için bu iki yöntem ile öznitelik seçimi gerçekleştirilebilmektedir.

İlk olarak SİA yöntemi ile seçilen öznitelikler irdelenmiştir. Bu yöntem ile verilen örnekleme en iyi temsil edecek öznitelikler seçilmeye çalışılmaktadır. İleri arama yöntemi şu şekilde çalışmaktadır. İlk olarak n tane bireysel elemandan oluşan başlangıç popülasyonu oluşturulur. Burada n giriş örneklemindeki öznitelik sayısıdır. Öznitelik alt kümeleri değerlendirilir ve en iyi k tanesi seçilir. Her bir k alt kümesi için kullanılmayan j öznitelik varsa, j adet öznitelik kopyası oluşturulur ve bir önceki adımda kullanılmayan tek bir öznitelik her bir adımda kümeye eklenir. Belirtilen p adet son döngüde performans artışı sağlandığı sürece ilk adımdan işleme devam edilir. Burada k başlangıçta belirlenen ve en az kaç öznitelik olacağını belirten bir sayıdır. Yine başlangıçta belirtilen bir diğer parametre ise p 'dir. Bu değer ise performans artışı son kaç döngüde artış sağlamazsa aramanın sonlandıracağını belirtir.

Diğer bir ifadeyle, sıralı ileri arama boş bir öznitelik seçimi ile başlar ve her bir adımda kullanılmayan bir öznitelik eklenir. Her eklenen öznitelik belirtilmiş bir değerlendirme fonksiyonu (sınıflandırıcı algoritma başarımı) ile değerlendirilir. Sadece en yüksek performans artışını sağlayan öznitelik seçime eklenir. Yeni tur değiştirilmiş öznitelik alt kümesi ile başlar. Belirlenen bir döngü sayısına ulaşınca ya da belirlenmiş son döngü sayısında performans artışı sağlanmadığında seçim işlemi sonlandırılır.

Burada k değeri 1 olarak belirlenmiştir. Yani en az 1 öznitelikten itibaren seçim işlemi başlatılmıştır. Sonlandırma kriterlerinden olan p değeri ise 5 olarak belirlenmiştir. Yani son 5 turda performans artışı sağlanmaz ise arama işlemi sonlandırılmaktadır. Seçim için döngü sayısı sınırı koyulmamıştır. SİA yöntemi ile 16 öznitelik seçilmiştir. Bu öznitelikler Tablo 18'de verilmiştir. Tabloda Seçim Durumu 1 olan öznitelikler SİA algoritması tarafından seçilmiş özniteliklerdir.

Tablo 18. SIA algoritması tarafından seçilen öznitelikler

Öznitelik	Seçim Durumu	Öznitelik	Seçim Durumu
SKEW0	0	STD2	0
KURT0	0	MEA2	1
RANG0	0	SKEW3	0
IQR0	0	KURT3	0
STD0	1	RANG3	0
MEA0	0	IQR3	1
SKEW1	1	STD3	0
KURT1	1	MEA3	1
RANG1	1	SKEW4	1
IQR1	1	KURT4	1
STD1	0	RANG4	0
MEA1	1	IQR4	1
SKEW2	0	STD4	0
KURT2	1	MEA4	1
RANG2	0	RRP	1
IQR2	0	RRA	1

SGA yönteminde ise SIA yönteminin tam tersi bir seçim stratejisi izlenir. Tüm özniteliklerden oluşan bir öznitelik kümesi ile seçime başlanır. Tüm öznitelik kümeleri değerlendirilir ve en iyi k tanesi alt küme belirlenir. Her bir k alt kümesi için eğer kullanılmayan j öznitelik var ise öznitelik kümesinin j adet kopyası oluşturulur ve her bir adımda önceki adımda kullanılan bir öznitelik silinir. Son p döngüde performans arttığı sürece işleme baştan devam edilir. Burada k ve p değerleri bir önceki yöntem ile aynı anlama gelmektedir.

Bir diğer ifadeyle, SGA özniteliklerin hepsini seçerek başlar ve her bir turda verilen öznitelik kümesinden bir tanesini çıkarır. Her öznitelik çıkarma işleminden sonra bir değerlendirme fonksiyonu tarafından performans tahmin edilir. Performansı en az düşüren öznitelik test kümesinden en son çıkarılır.

Bir önceki arama yönteminde kullanılan parametre setine benzer şekilde k değeri için 1 ve p değeri için 5 değerleri kullanılmıştır. SGA yöntemi ile 24 öznitelik seçilmiştir. Bu öznitelikler Tablo 19’da verilmiştir. Tabloda Seçim Durumu 1 olan öznitelikler geri arama algoritması tarafından seçilmiş özniteliklerdir.

Tablo 19. SGA tarafından seçilen öznitelikler

Öznitelik	Seçim Durumu	Öznitelik	Seçim Durumu
SKEW0	1	STD2	1
KURT0	1	MEA2	1
RANG0	0	SKEW3	1
IQR0	1	KURT3	1
STD0	1	RANG3	0
MEA0	0	IQR3	1
SKEW1	1	STD3	0
KURT1	1	MEA3	1
RANG1	1	SKEW4	0
IQR1	1	KURT4	1
STD1	0	RANG4	1
MEA1	1	IQR4	1
SKEW2	1	STD4	1
KURT2	1	MEA4	1
RANG2	0	RRP	1
IQR2	0	RRA	1

2.6.4. Sınıflandırma

Bir vuruluk EKG işaretinden oluşan zaman serisinden elde edilen 32 öznitelik, bu öznitelikleri temsil edecek seçilmiş alt kümeleri ve hesaplanmış daha düşük boyutlu versiyonları sınıflandırma algoritmaları tarafından sınıflandırılmış ve sonuçları karşılaştırmalı olarak irdelenmiştir. Başlık 1.10'da sınıflandırma başarımlarının değerlendirilmesi için başarımlar ölçütleri verilmiştir. Tablo 4'te verilen karşıtlık matrisi iki sınıflı sınıflandırma için geçerlidir. Sınıflandırma sistemi çok sınıflı olduğunda karşıtlık matrisi yine test sonuçlarından hesaplanabilir. Ancak sınıflandırma başarımlar ölçütlerinin hesaplanması biraz farklılık gösterir. Çalışmanın bu bölümünde 9 farklı vuru türü sınıflandırılacağı için karşıtlık matrisi Tablo 20'de gösterildiği gibi olacaktır.

Tablo 20. Çok sınıflı sınıflandırma için karşıtlık matrisi

		Sınıflandırma sistemi çıktısı									Duyarlılık(D)
		N	/	E	F	L	R	V	a	f	
Gerçek Durum	N	TP_N	FN_N	FN_N	FN_N	FN_N	FN_N	FN_N	FN_N	FN_N	D_N
	/	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	$D_{/}$
	E	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	D_E
	F	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	D_F
	L	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	D_L
	R	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	D_R
	V	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	D_V
	a	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	D_a
	f	FP_N	TN	TN	TN	TN	TN	TN	TN	TN	D_f
Kesinlik (K)	K_N	$K_{/}$	K_E	K_F	K_L	K_R	K_V	K_a	K_f	Genel Doğruluk	

Tabloda normal (N) sınıfı için sınıflandırma karşıtlık tablosu verilmiştir. Her bir sınıf için benzer şekilde tablodaki değerlere göre başarımların hesaplanacaktır.

Burada;

TP – Doğru Pozitif: Sistem tarafından bir aritmi türünde doğru sınıflandırılmış vuru sayısı

TN – Doğru Negatif: Sistem tarafından diğer aritmi türlerinde doğru sınıflandırılmış vuru sayısı

FP – Yanlış Pozitif: Sistem tarafından yanlış olarak sınıflandırılan vuru sayısı

FN – Yanlış Negatif: Sistem tarafından yanlış sınıflandırılan diğer vuruların sayısı

Çok sınıflı sınıflandırmada doğruluk değeri tek olmasına rağmen duyarlılık, özgünlük ve kesinlik değerleri her sınıfa göre ayrı ayrı hesaplanmaktadır. Eşitlik (76)-(79) çok sınıflı sınıflandırma için performans kriterlerini göstermektedir.

$$Duyarlılık_i = \frac{TP_i}{TP_i + \sum FN_i} \quad (76)$$

$$Özgünlük_i = \frac{TN_i}{TP_i + \sum FP_i} \quad (77)$$

$$Kesinlik_i = \frac{TP_i}{TP_i + \sum FP_i} \quad (78)$$

$$Genel Doğruluk = \frac{\sum_{i=1}^9 TP_i}{\text{Toplam Vuru Sayısı}} \quad (79)$$

Burada alt indis i ilgili sınıf türünü göstermektedir. Her sınıfa ait kesinlik ve duyarlılık kriterlerini incelemek yerine bunların ortalaması alınarak sınıflandırıcı performansı değerlendirilebilir. Ortalama kesinlik ve ortalama duyarlılık değerleri eşitlik (80) ve (81)'de verilmiştir.

$$\text{Ortalama Duyarlılık (OD)} = \frac{\sum_{i=1}^9 Duyarlilik_i}{9} \quad (80)$$

$$\text{Ortalama Kesinlik (OK)} = \frac{\sum_{i=1}^9 Kesinlik_i}{9} \quad (81)$$

Yukarıda belirtilen başarımların kriterleri testlerde kullanılan tüm sınıflandırıcılar için hesaplanmış ve karşılaştırma işlemlerinde kullanılmıştır. Testlerde sınıflandırıcı olarak YSA, K-NN, DVM ve KA kullanılmıştır. Sınıflandırma işlemlerinde 10-kere çapraz doğrulama yöntemi kullanılmıştır. 10-kere çapraz doğrulama işleminde ilk olarak veri kümesi aynı boyutlu on ayrı alt kümeyle bölünür. Sırasıyla her seferinde bir tane alt küme test için ayrılırken geri kalan 9 alt kümesi ise sistemin eğitimi için kullanılır. On tekrarın

sonunda elde edilen performans ölçütlerinin ortalaması alınır. Böylece veri setindeki her bir alt küme bir kez tahmin edilmiş olur. Eşitlik (82)-(84)'te 10-kere çapraz doğrulama doğruluğu ($\mathcal{C}DD$), ortalama duyarlılığı ($\mathcal{C}DOD$) ve ortalama kesinliği ($\mathcal{C}DOK$) gösterilmiştir.

$$\mathcal{C}DD = \frac{1}{k} \sum_{i=1}^k D_i \quad (82)$$

$$\mathcal{C}DOD = \frac{1}{k} \sum_{i=1}^k OD_i \quad (83)$$

$$\mathcal{C}DOK = \frac{1}{k} \sum_{i=1}^k OK_i \quad (84)$$

Burada D_i , OD_i ve OK_i sırasıyla her bir denemede elde edilen doğruluk, ortalama duyarlılık ve ortalama kesinlik değerlerini, i deneme numarasını ve k çapraz doğrulama sayısını ($k=10$) göstermektedir.

YSA ile Sınıflandırma

Testlerde ilk olarak çok katmanlı YSA mimarisi kullanılmıştır. Giriş katman öznelik boyutu ile aynı olacak şekilde (32 veya azaltılmış boyut sayısı kadar), çıkış katmanı ise sınıf sayısı ile aynı olacak şekilde 9 olarak belirlenmiştir. Tasarlanan YSA mimarisinde bir adet gizli katman olup gizli katmanın düğüm sayısı 20 olarak belirlenmiştir. Gizli katman düğüm sayısını belirlemek için 2-20 arasındaki sayılar için sınıflandırma başarımları irdelenmiş ve testler sonucunda elde edilen başarımlar incelenmiştir. Tablo 21'de 7 öznelik grubu için elde edilen sonuçlar verilmiştir.

Tablo 21. Öznitelik grupları için YSA gizli katman düğüm sayılarının sınıflandırma başarımına etkisi

Gizli Katman Düğüm sayısı	Tüm Öznitelikler	TBA	BBA	ÖH	İleri Arama	Geri Arama	GA
2	87,56%	86,89%	84,41%	53,37%	85,99%	86,74%	86,10%
4	97,39%	97,10%	96,50%	68,03%	96,41%	97,14%	96,60%
6	98,24%	97,97%	97,55%	74,64%	97,94%	98,01%	97,59%
8	98,26%	98,21%	98,15%	77,24%	98,20%	98,42%	97,97%
10	98,42%	98,59%	98,12%	79,41%	98,32%	98,62%	98,24%
12	98,67%	98,61%	98,32%	81,91%	98,54%	98,64%	98,53%
14	98,77%	98,70%	98,39%	82,65%	98,62%	98,65%	98,70%
16	98,73%	98,69%	98,61%	82,97%	98,51%	98,78%	98,65%
18	98,65%	98,64%	98,46%	84,36%	98,92%	98,65%	98,62%
20	98,73%	98,89%	98,62%	85,68%	98,75%	98,81%	98,94%

Tabloda da görüleceği gibi tüm öznitelikler kullanıldığında gizli katman düğüm sayısı 14 en iyi sonucu vermişken öznitelik boyutu azaltıldığında gizli katman düğüm sayısı 20 civarında en iyi sonucu vermiştir. Çalışmada kullanılan YSA geri yayılım algoritması kullanılarak eğitilmiştir. Eğitim aşamasında tekrar sayısı parametresi 500 ve öğrenme katsayısı 0,3 olarak belirlenmiştir. Hata eşiği parametresi ise 0,00001 olarak ayarlanmıştır.

K-NN ile Sınıflandırma

Çalışmada kullanılan bir diğer sınıflandırıcı K-NN sınıflandırıcıdır. K-NN sınıflandırıcı ile sınıflandırma yapılırken testlerde en iyi sonucu verecek K değerinin hesaplanması için ızgara arama yöntemi kullanılmıştır. Arama işlemi sonucunda K değeri bir olarak hesaplanmıştır. Arama aralığındaki tüm K değerleri yüksek başarımler elde etmiştir. K-NN sınıflandırıcıda sonuçları elde etmek için öklit uzaklık fonksiyonu kullanılmıştır. K-NN algoritmasının elde ettiği başarımler ve hesaplama süresinin kısa olmasından dolayı parametre optimizasyonu işlemlerinde değerlendirme fonksiyonu olarak K-NN sınıflandırıcısı kullanılmıştır.

Tablo 22. Öznitelik grupları için K-NN algoritmasında K değerinin başarıma etkisi

K	Tüm Öznitelikler	TBA	BBA	ÖH	İleri Arama	Geri Arama	GA
1	98,86%	98,86%	99,11%	98,26%	99,37%	99,29%	99,30%
3	98,73%	98,73%	98,80%	97,82%	99,19%	99,15%	99,13%
5	98,53%	98,53%	98,62%	97,26%	98,96%	98,96%	98,83%
7	98,43%	98,43%	98,53%	96,95%	98,83%	98,84%	98,67%
9	98,21%	98,21%	98,29%	96,53%	98,58%	98,58%	98,43%
11	98,07%	98,07%	98,23%	96,00%	98,29%	98,27%	98,20%
13	97,97%	97,97%	98,02%	95,65%	98,26%	98,13%	98,16%
15	97,80%	97,80%	97,88%	95,33%	98,12%	98,04%	98,01%

Tablo 22’de tüm öznitelik grupları için K-NN sınıflandırıcısı ile sınıflandırırken K değerinin sınıflandırma başarımına etkisi gösterilmiştir. Burada 1-15 aralığındaki tüm tek sayılar K değeri olarak kullanılmış ve elde ettikleri başarımler listelenmiştir.

DVM ile Sınıflandırma

DVM ile sınıflandırma sonuçlarını elde etmek için öncelikle DVM parametrelerinin belirlenmesi gerekmektedir. Burada DVM parametreleri ızgara arama yöntemi kullanılarak belirlenmiştir. Parametre optimizasyonu için DVM’ye ait çekirdek fonksiyonları, karmaşık sabit DVM parametresi C ve Gamma parametreleri en iyi sonucu verecek şekilde optimize edilmeye çalışılmıştır.

Tablo 23’te çekirdek fonksiyonu olarak lineer seçildiğinde C parametresinin değişimi ile elde edilmiş sınıflandırma başarımı gösterilmiştir.

Tablo 23. Lineer çekirdek fonksiyonu için C parametresinin başarıma etkisi

C	Doğruluk
0	96,85%
0,1	98,64%
0,2	98,59%
0,3	98,64%
0,4	98,65%
0,5	98,61%
0,6	98,54%
0,7	98,67%
0,8	98,50%
0,9	98,70%
1	98,64%
2	98,64%
10	98,70%
100	98,48%
1000	98,39%

Çekirdek fonksiyonu olarak polinom seçildiğinde optimize edilmesi gereken parametrelere derece eklenmektedir. Derece, C ve Gamma parametreleri polinom çekirdek fonksiyonu için birlikte aranmış ve $C=1$, Gamma=1 ve derece=3 değerleri en iyi sonucu veren parametreler olmuştur. Aramalarda derece değeri 2-4 arasında değişen değerlerde test edilmiştir. Tablo 24'te çekirdek fonksiyonu seçimlerinin sınıflandırma başarıma etkisi gösterilmiştir.

Tablo 24. Seçilen çekirdek fonksiyonlarının başarıma etkisi

Çekirdek Fonksiyonu	Doğruluk
Lineer	98,70%
Polinom	99,05%
RBF	97,53%
Sigmoid	14,07%

KA ile Sınıflandırma

Çalışmada elde edilen diğer sonuçlarla karşılaştırmak için KA sınıflandırma yöntemi de testlerde kullanılmış ve elde ettiği sınıflandırma başarıma değerlendirilmiştir. KA ile

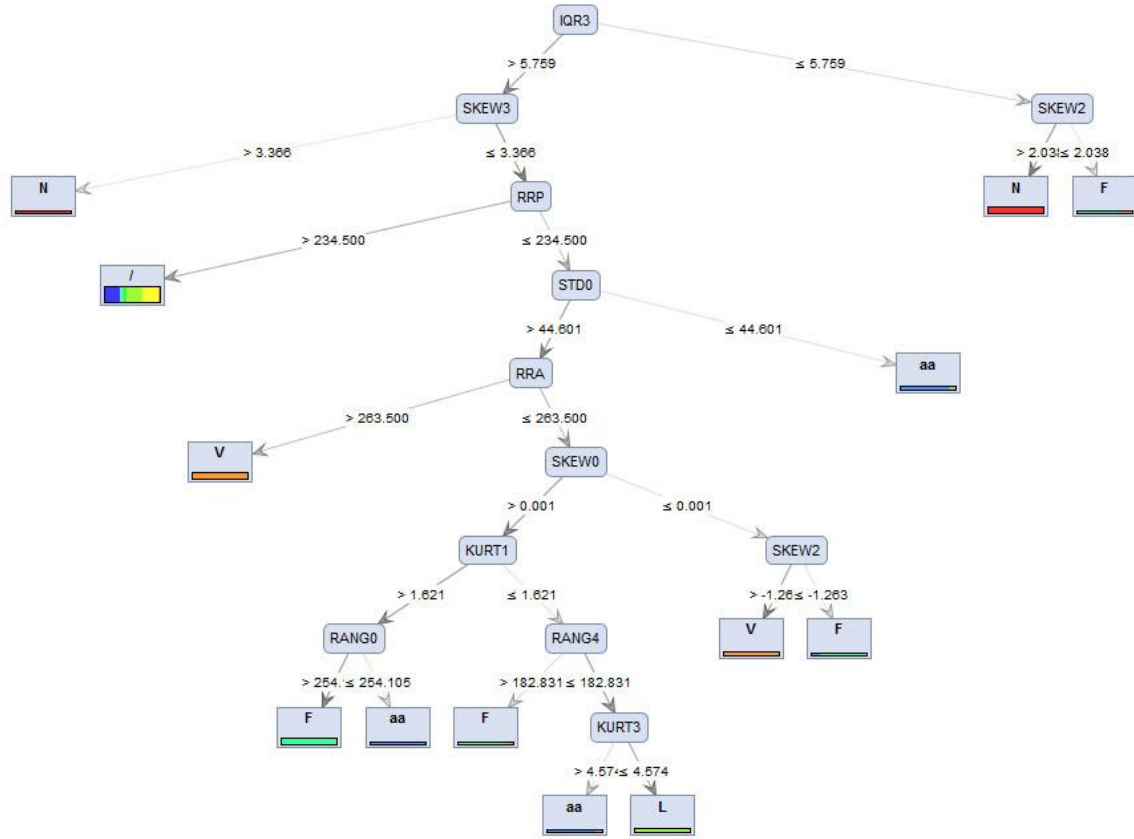
verinin sınıflandırılmasında modelin yorumlanmasının kolay olması ve gösteriminin anlamlı olması diğer yaklaşımlar ile karşılaştırıldığında daha avantajlıdır.

KA ile sınıflandırma yapılırken önceki yöntemlere benzer şekilde parametre optimizasyonu için ızgara arama yöntemi ile en iyi sonucu verecek parametre değerleri araştırılmıştır. Bu amaçla KA için maksimum derinlik (MD) ve minimum yaprak sayısı (MYS) parametreleri belirlenmiş aralıklarda değiştirilerek en iyi sonuç elde edilmeye çalışılmıştır. MD parametresi için 10-30 arasındaki değerler, MYS parametresi için ise 2-5 arasındaki değerler hesaplamalarda teker teker kullanılmıştır. Tablo 25'te parametre optimizasyonu aşamasında elde edilmiş sınıflandırma başarımı gösterilmiştir.

Tablo 25. KA sınıflandırmada parametre optimizasyonu

MD	MYS			
	2	3	4	5
10	73,62%	69,81%	70,45%	69,97%
11	74,19%	70,63%	70,80%	71,75%
12	67,00%	74,28%	76,06%	66,96%
13	70,33%	70,59%	69,21%	70,69%
14	70,09%	74,26%	73,68%	66,85%
15	70,59%	74,04%	74,36%	69,79%
16	73,96%	70,58%	70,64%	67,51%
17	69,97%	66,73%	73,66%	70,53%
18	73,32%	73,94%	74,08%	70,11%
19	70,29%	69,79%	70,22%	70,12%
20	74,16%	66,90%	70,16%	71,07%
21	70,29%	70,21%	69,98%	69,18%
22	70,52%	70,78%	74,01%	70,35%
23	69,36%	74,55%	74,20%	69,99%
24	70,14%	74,11%	67,83%	74,51%
25	73,70%	70,62%	74,26%	66,89%
26	73,23%	70,07%	70,92%	69,87%
27	70,27%	74,28%	70,13%	70,03%
28	70,72%	70,43%	74,34%	67,08%
29	66,08%	73,85%	69,34%	65,42%
30	73,69%	70,58%	69,99%	70,07%

KA kullanılarak sınıflandırma işlemi gerçekleştirilirken oluşturulmuş bir karar ağacı Şekil 65'te gösterilmiştir.



Şekil 65. EKG vurularını sınıflandıran örnek Karar Ağacı

2.6.5. Sınıflandırma Sonuçları

Deneylerde kullanılan sınıflandırma algoritmalarının başarımlarının değerlendirilmesi için 10-kez çapraz doğrulama sonucunda elde edilmiş *doğruluk*, *ortalama duyarlılık* ve *ortalama kesinlik* değerleri hesaplanmıştır. Tablo 26'da sınıflandırıcılara giriş verisi olarak hesaplanmış 32 özniteliğin kullanılması ile elde edilen sınıflandırma sonuçları gösterilmiştir.

Tablo 26. Giriş verisi olarak 32 istatistiksel öznelik için sınıflandırma sonuçları (%)

	ÇDD	ÇDOD	ÇDOK
K-NN	98,86%	97,93%	97,79%
YSA	98,86%	98,39%	97,26%
DVM	98,92%	98,43%	97,58%
KA	70,94%	63,93%	67,58%

Tablo 27 TBA ile hesaplanan öznelikleri giriş vektörü olarak kullanan sınıflandırıcıların elde ettikleri başarımları göstermektedir.

Tablo 27. Giriş verisi olarak TB'ler için sınıflandırma sonuçları (%)

	ÇDD	ÇDOD	ÇDOK
K-NN	98,86%	97,93%	97,79%
YSA	98,86%	97,96%	97,80%
DVM	98,72%	98,41%	97,67%
KA	85,50%	84,14%	85,80%

Tablo 28 BBA ile hesaplanan öznelikleri giriş vektörü olarak kullanan sınıflandırıcıların elde ettikleri başarımları göstermektedir.

Tablo 28. Giriş verisi olarak BB'ler için sınıflandırma sonuçları (%)

	ÇDD	ÇDOD	ÇDOK
K-NN	99,11%	98,49%	98,26%
YSA	98,59%	98,32%	96,81%
DVM	98,50%	97,97%	97,01%
KA	91,39%	85,86%	91,14%

Tablo 29 ÖH ile hesaplanan öznelikleri giriş vektörü olarak kullanan sınıflandırıcıların elde ettikleri başarımları göstermektedir.

Tablo 29. Giriş verisi olarak ÖH için sınıflandırma sonuçları (%)

	ÇDD	ÇDOD	ÇDOK
K-NN	98,26%	97,50%	97,43%
YSA	84,65%	82,64%	82,05%
DVM	63,42%	45,73%	44,45%
KA	86,78%	83,99%	85,86%

Tablo 30 GA ile seçilen öznitelikleri giriş vektörü olarak kullanan sınıflandırıcıların elde ettikleri başarımları göstermektedir.

Tablo 30. Giriş verisi olarak GA ile seçilen öznitelikler için sınıflandırma sonuçları (%)

	ÇDD	ÇDOD	ÇDOK
K-NN	99,30%	98,84%	98,40%
YSA	98,65%	97,91%	96,95%
DVM	98,80%	98,13%	97,19%
KA	94,65%	81,75%	80,42%

Tablo 31 İleri arama ile seçilen öznitelikleri giriş vektörü olarak kullanan sınıflandırıcıların elde ettikleri başarımları göstermektedir.

Tablo 31. Giriş verisi olarak SİA ile seçilen öznitelikler için sınıflandırma sonuçları (%)

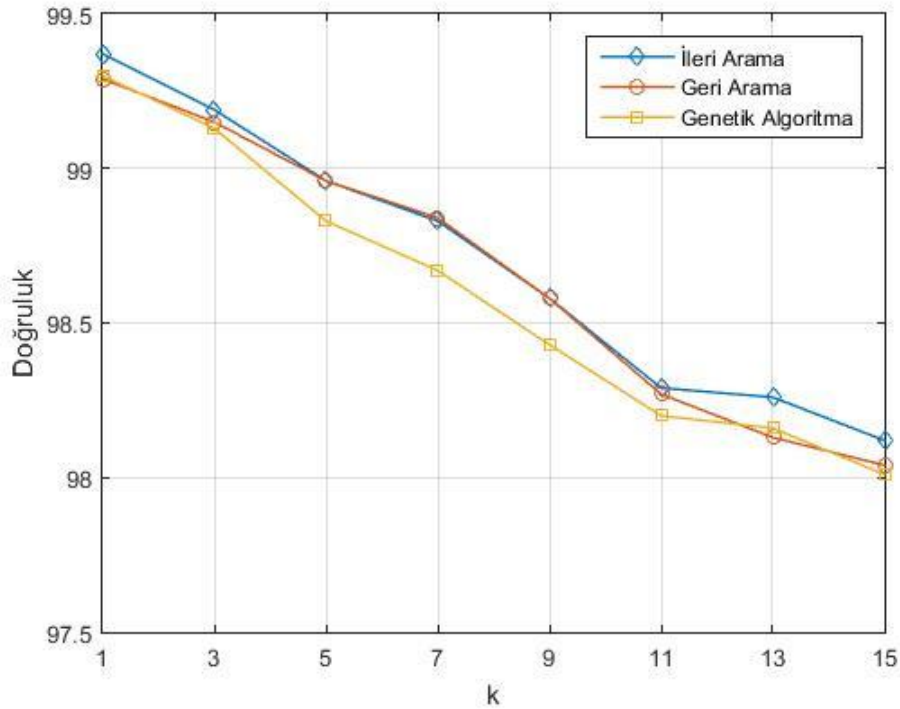
	ÇDD	ÇDOD	ÇDOK
K-NN	99,37%	99,01%	98,69%
YSA	98,88%	98,40%	97,73%
DVM	98,83%	98,12%	97,29%
KA	95,20%	86,31%	84,31%

Tablo 32 Geri arama ile seçilen öznitelikleri giriş vektörü olarak kullanan sınıflandırıcıların elde ettikleri başarımları göstermektedir.

Tablo 32. Giriş verisi olarak SGA ile seçilen öznitelikler için sınıflandırma sonuçları (%)

	ÇDD	ÇDOD	ÇDOK
K-NN	99,29%	98,95%	98,43%
YSA	98,99%	98,56%	97,78%
DVM	98,80%	98,13%	97,40%
KA	70,48%	61,44%	64,07%

Deneyleerde elde edilen sonuçlardan görüldüğü gibi K-NN sınıflandırıcı tüm giriş vektörlerinde başarılı sonuçlar elde etmektedir. K-NN sınıflandırıcının elde ettiği başarımdan dolayı sınıflandırıcıya ait tek parametre olan K değerini hesaplamak için parametre optimizasyonu yapılmıştır. En yüksek başarıyı K=1 ve giriş vektörü olarak SİA ile seçilmiş özniteliklerin kullanıldığı K-NN sınıflandırma şeması elde etmiştir. K değerleri irdelenmiş ve sonuçların kabul edilebilir düzeyde birbirine yakın olduğu görülmüştür. En yüksek başarımlar istatistiksel özniteliklerden seçim algoritmaları ile seçilen (SİA, SGA ve GA) öznitelikler ile elde edilmiştir. Bu öznitelik grubu için belirtilen aralıktaki tüm k değerleri kabul edilebilir düzeyde başarı (>%98) elde etmiştir.



Şekil 66. K sayısının sınıflandırma başarımına etkisi

Şekil 66’te üç öznitelik seçim algoritması ile seçilmiş olan öznitelik grubu için K-NN ile sınıflandırmada K parametresinin sınıflandırma başarımına etkisi gösterilmiştir. K değeri arttıkça sınıflandırma başarımının az da olsa düştüğü gözlenmektedir. Ancak bu düşüş sınırlı kalmıştır.

Özellikle k değerinin 3 ve 5 olduğu durumlar için %99 civarı başarımlar elde ettiği görülmüştür. SİA ile elde edilmiş öznitelik vektörünün aralıktaki tüm K değerlerinde en yüksek sınıflandırma başarımı elde ettiği gösterilmiştir.

Tablo 33. 10-kez çapraz doğrulama için her adımda elde edilen sınıflandırma başarımı

Döngü	Doğruluk	OD	OK
1	99,21%	99,33%	98,64%
2	99,37%	99,53%	98,03%
3	99,52%	98,98%	98,50%
4	99,37%	98,90%	98,63%
5	99,37%	98,63%	99,50%
6	98,89%	98,48%	97,18%
7	99,53%	98,64%	99,64%
8	99,52%	99,01%	99,64%
9	99,37%	99,56%	98,16%
10	99,53%	99,04%	99,01%
Ortalama	99,37%	99,01%	98,69%

Çalışmadaki tüm deneylerde 10-kez çapraz doğrulama yapıldığı daha önceki örneklerde bahsedilmişti. K-NN sınıflandırıcı için giriş vektörü olarak SİA ile elde edilmiş öznitelik vektörünün kullanıldığı testlerde elde edilen sonuçlar için her bir adımdaki çapraz doğrulama sonuçları Tablo 33’te verilmiştir.

10-kez çapraz doğrulamada elde edilen sonuç her bir denemede elde edilen sonuçların ortalamasıydı. Tabloya dikkat edilecek olursa sınıflandırma başarımı olarak kabul edilmiş başarımdan daha yüksek başarımların da çoğu döngüde elde edildiği görülebilir. Buradan sınıflandırma başarımı olarak 10-kez çapraz doğrulama yönteminin seçilmesinin doğru bir karar olduğu görülmektedir. Klasik yöntemlerde bu denemelerden her biri sonuç olarak kabul edilebileceği için daha yüksek veya daha düşük sonuçlar elde edilmiş olabilir.

2.6.6. Büyük Veri Testi Sonuçları

Çalışmada en başarılı sonuçları SİA ile seçilen öznelikleri kullanan K-NN sınıflandırma algoritması elde etmişti. Önerilen bu model kullanılarak veritabanında bulunan 9 sınıfa ait toplam 100457 vuru sınıflandırılmış ve sonuçlar değerlendirilmiştir. Testte kullanılan vuruların dağılımı ve toplam içerisindeki oranı Tablo 34’te gösterilmiştir.

Tablo 34. Büyük veri testi için vuruların sınıflara dağılımı ve oranları

Sınıf	Sayı	Oran
N	73326	72,992%
L	8066	8,029%
R	7248	7,215%
V	6885	6,854%
/	3614	3,598%
F	802	0,798%
f	260	0,259%
a	150	0,149%
E	106	0,106%
Toplam	100457	

Tabloda bulunan veriler 10-kez çapraz doğrulama ile sınıflandırıldığında sırasıyla ÇDD, ÇDOD ve ÇDOK sonuçları, %99,08, %92,03, %95,23 olarak hesaplanmıştır. Burada dikkat edilecek olursa ÇDOK ve ÇDOK sonuçları genel doğruluk sonucuna göre daha düşük çıkmıştır. Bunun sebebi ise sınıflardaki örnek dağılımının arasında çok fark olmasıdır. Sonuçların daha anlaşılabilir olması için ÇDOK ve ÇDOK değerleri hesaplanırken sınıf dağılımları göz önünde bulundurulduğunda (ağırlıklı ortalama) sırasıyla %99,08 ve %99,07 sınıflandırma başarımlarına ulaşılmıştır. Tablo 35’te sınıflandırma sonuçlarının hesaplandığı karşıtlık matrisi gösterilmiştir.

Tablo 35. Büyük veri sınıflandırma karşıtlık matrisi

		Gerçek Durum									Kes.
		N	V	/	F	L	a	R	E	f	
Sınıflandırıcı Tahmini	N	73026	158	1	131	45	34	49	4	14	99,41%
	V	96	6635	0	55	6	6	0	0	0	97,60%
	/	1	2	3610	0	0	0	0	0	4	99,81%
	F	64	58	0	612	0	0	0	0	0	83,38%
	L	69	19	0	4	8008	0	1	2	1	98,82%
	a	11	8	0	0	0	110	0	0	0	85,27%
	R	51	5	0	0	5	0	7197	3	0	99,12%
	E	0	0	0	0	2	0	1	97	0	97,00%
	f	8	0	3	0	0	0	0	0	241	95,63%
Duy.	99,59%	96,37%	99,89%	76,31%	99,28%	73,33%	99,30%	91,51%	92,69%	99,08%	

2.6.7. Öznitelik Azaltma ve Sınıflandırma Hesaplama Süreleri

Deneysel olarak kullanılan sınıflandırıcı ve giriş vektörlerine ait 10-kez çapraz doğrulama için test süreleri (sn.) ölçülmüş ve Tablo 36’da verilmiştir. Hesaplanan bu sürelerin giriş vektör boyutuna ve kullanılan sınıflandırıcı mimarisine göre değiştiği gözlenmektedir. En küçük boyutlu (3) öznitelik vektörü olan ÖH sınıflandırıcılarda en hızlı işlem gören öznitelik grubu olmuştur. Sınıflandırıcı temelinde YSA hesaplama yükünden dolayı en yavaş sınıflandırıcı olarak sonuç vermektedir.

Tablo 36. Farklı giriş vektörleri için sınıflandırma algoritmalarının hesaplama süreleri (sn)

	İstatistiksel Öznitelikler	TBA	BBA	ÖH	GA	SİA	SGA
K-NN	3,91	3,452	3,216	1,491	3,647	4,412	5,648
YSA	741,464	485,161	370,835	241,879	432,13	393,15	503,439
DVM	3,958	2,944	2,598	14,259	2,762	3,047	2,525
KA	6,053	3,957	2,268	0,774	3,987	3,613	4,359

Öznitelik azaltma algoritmalarının hesaplama süreleri ölçülmüş ve sonuçları (sn.) Tablo 37’de gösterilmiştir. TBA ve BBA içerisinde herhangi bir doğrulama işlemi olmayan matematiksel yöntemler olduğu için tezde kullanılan veri seti için en hızlı hesaplamaları gerçekleştirmiştir. ÖH yapısı gereği YSA alt yapısına benzerliğinden dolayı en çok zaman

alan öznitelik azaltma yöntemi olmuştur. GA, SİA ve SGA yöntemleri içerisinde 10-kez çapraz doğrulama yaparak seçilen öznitelik alt kümesini değerlendiren uzun arama adımları içerdiğinden TBA ve BBA yöntemine göre daha uzun hesaplama süreleri olmuştur.

Tablo 37. Öznitelik azaltma algoritmalarının hesaplama süreleri (sn.)

TBA	BBA	ÖH	GA	İleri Arama	Geri Arama
0,161	0,21	2232,082	691,99	871,08	1066,131

Burada dikkat edilmesi gereken nokta TBA, BBA ve ÖH gibi yöntemler kullanıldığında her sınıflandırma işleminde bu öznitelik azaltma yöntemlerinin tekrar hesaplanacak olmasıdır. Ancak GA, SİA ve SGA yöntemleri öznitelik seçim yöntemleri olduğu için test aşamasında seçilen öznitelikler geriye doğru ilk hesaplama adımına dönülüp o aşamada seçim dışı kalan özniteliklerin hesaplanması engellenerek hazırlanacak olan sistemler için performans iyileştirmesi sağlanmış olur.

3. SONUÇLAR

Bu tezde yapılan çalışmaları iki ana başlıkta toplamak sonuçların değerlendirilmesi açısından yerinde olacaktır. Çalışmada ilk olarak EKK virusunun doğru şekilde algılanması için sınıflandırma algoritmaları ve boyut azaltma algoritmalarının performansları incelenmiştir. Yapılan deneyler sonucunda EKK virusunun sınıflandırılması için en iyi başarıyı elde edecek sınıflandırma şeması belirlenmiştir. İkinci bölümde ise 9 ayrı vuru türünü sınıflandırmak için sınıflandırma ve boyut indirgeme şemalarının performansları değerlendirilmiş, en iyi performansı verecek yaklaşımlar belirlenmiştir.

3.1. EKK Virusunun Sınıflandırılması ile İlgili Sonuçlar ve Değerlendirme

EKK virusunun sınıflandırılması için işarete ait morfolojik herhangi bir öznelik çıkarılmamış olup bir vuruluk işarete ait genlik değerleri (zaman serisi) kullanılmıştır. QRS kompleksini ortalayacak şekilde 200 örnekten oluşan zaman serisi sınıflandırma algoritmaları tarafından test edilmiş ve karşılaştırmalarda kullanılmak üzere TBA, BBA, ÖH ve GA yöntemleri ile boyutu indirgenmiştir. TBA ile boyut indirgenildiğinde en iyi sonucu veren TB sayısı olarak 17 değeri ızgara arama sonucunda elde edilmiştir. Boyut indirgeme amacı ile BBA işarete uygulandığında BB sayısı 10 olarak hesaplanmıştır. Çalışmada kullanılan bir diğer boyut indirmeğe yöntemi ÖH'dır. ÖH 200 örnekten oluşan giriş uzayını iki boyuta indirmek için kullanılmıştır.

Diğer boyut indirgeme yöntemlerinden farklı olarak GA var olan özneliklerden veriyi en iyi temsil edecek olanların seçimi için kullanılmıştır. 200 örnekten oluşan giriş verisinin daha düşük örneklerle temsili için daha değerli özneliklerin seçimi için GA kullanılmıştır. GA ile seçim yapılırken farklı üst sınırlarda arama işlemleri gerçekleştirilmiştir. 25, 50 ve 100 değerleri maksimum öznelik boyutu olarak belirlenmiş ve bu sınırlarda öznelikler belirlenmeye çalışılmıştır. Bu değerlere karşılık gelen 25, 50 ve 95 adet öznelik GA tarafından seçilmiştir.

Deneylerde YSA, K-NN, DVM ve KA sınıflandırıcıları sonuçları karşılaştırılmak üzere farklı sınıflandırma şemalarında testlere tabi tutulmuştur. Deneylerde ilk olarak çok

katmanlı YSA mimarisi kullanılmıştır. Tasarlanan YSA mimarisinde bir adet gizli katman olup gizli katmandaki düğüm sayısı parametre optimizasyonu ile 10 olarak belirlenmiştir. K-NN çalışmada EKK vurularının sınıflandırılması için kullanılan bir diğer sınıflandırıcıdır. K-NN sınıflandırıcısında k değerinin hesaplanması için ızgara arama yöntemi kullanılarak en iyi sonucu verecek k değerleri araştırılmıştır. DVM sınıflandırıcısı için C, Gamma ve çekirdek fonksiyonları için ızgara arama yöntemi ile parametre optimizasyonu yapılmış ve parametre değerleri belirlenmiştir.

Tablo 38. Çalışmada kullanılan sınıflandırıcılar ve yayınlanmış EKK sınıflandırıcılarının yazarları tarafından raporlanan performans metrikleri (%)

Araştırmacılar	Kullanılan Sınıflandırıcı	Özgünlük	Duyarlılık	Doğruluk	İşaret Sayısı
De Oliveira vd. [184]	Bayes Ağları	99,86	95,09	-	11
Ittatur vd. [185]	Basit karar kuralı	99,55	91,05	-	26
Bortolan vd. [186]	YSA, K-NN, DA, FL	98,7	91,3	-	48
Li vd. [187]	Şablon eşleştirme	-	93,1	98,2	22
Ribeiro vd. [188]	DVM	98,28	89,39	-	-
Foo vd. [189]	YSA	-	-	92,2	4
Inan vd. [190]	YSA	-	98,33	95,16	40
Jenny vd. [191]	k-ortalama, Bulanık c-ortalama	81,1	80,1	80,94	-
Talbi vd. [192]	Kesirli lineer tahmin	95	85	95	43
Tezdeki Yaklaşım	GA+K-NN	99,91	99,46	99,69	43
	GA+YSA	98,14	98,91	98,53	
	GA+DVM	99,31	96,89	98,10	
	GA+KA	98,57	96,77	97,67	

Test edilen sınıflandırıcıların çoğu yüksek başarımlar elde etmiştir. Özellikle K-NN sınıflandırıcı en yüksek başarımlar olan %99,69 değerini GA ile seçilmiş öznitelikleri kullanarak elde etmiştir. KA EKK vurularının sınıflandırılmasında en başarısız sınıflandırıcı olmuştur. KA ve DVM sınıflandırıcıları ÖH özniteliklerini kullanarak sırasıyla %81,39 ve %77,54 sınıflandırma başarımları elde ederek en düşük sonuçları elde etmiştir. Sınıflandırma süreleri dikkate alındığında yine boyutu indirgenmiş vektörleri giriş vektörü olarak kullanan K-NN sınıflandırıcı en düşük sınıflandırma sürelerini elde etmiştir. Deneylede kullanılan

tüm sınıflandırıcılar boyut indirgeme ile kayda değer hızlanma elde etmişlerdir. YSA mimarisi gereği diğer sınıflandırıcılara göre daha uzun hesaplama süreleri kullanmıştır.

Tablo 38’de EKG vurularının sınıflandırılması için önerilen yöntemin elde ettiği sonuç ile literatürde aynı başlıkta yapılmış çalışmalara ait sonuçların karşılaştırılması verilmektedir.

Doğruluk, özgünlük ve duyarlılık kriterleri sınıflandırma algoritmalarının başarımını karşılaştırmak için hesaplanmıştır. Önerilen yaklaşımda diğer algoritmalar ile karşılaştırmalı olarak gösterilmiştir ki GA ile seçilmiş öznitelikleri kullanan K-NN sınıflandırıcı %99,69 doğruluk oranı ile en yüksek başarımı elde etmiştir. Bu bağlamda önerilen yöntem yapılan diğer çalışmalarda kullanılan yöntemlerden daha yüksek başarımlar elde etmiştir. Şu anki çalışmaların çoğu ([184,185,187,189]) veritabanındaki verilerin bir bölümünü kullanmıştır. Bu çalışmada veritabanında bulunan EKK vurularının tamamına yakını kullanılmıştır. De Oliveira vd. sınıflandırma için 947 EKK vurusu kullanmış ve bunların %80’ini eğitim %20’sini ise test için kullanmıştır. Çalışmada çapraz doğrulama kullanılıp kullanılmadığı gibi deneysel uygulama ile yeterli bilgi verilmemiştir. Vuruların veritabanındaki hangi kayıtlardan alındığına dair bilgi yoktur [184]. Bir diğer çalışmada Ittatur ve arkadaşları yöntemlerini 26 kayıt üzerinde test etmiş ve kalp pili, kalp blokajı ve atriyal çarpınım içeren bazı kayıtları testlerin dışında tutmuşlardır [185]. Diğer yandan Bortolan vd. çalışmalarında veritabanındaki tüm kayıtları kullanmış olmasına rağmen testlerde kullandığı vuru sayısı sınırlı kalmıştır. Eğitim için kullandığı veri seti (260 vuru genel set ve 76 vuru yerel set) boyutu çok küçük tutulmuştur. Yazarlar en iyi sınıflandırma başarımını ise DA kullanarak genel set diye tanımladıkları veri setinde %88,5 doğruluk ve K-NN kullanarak yerel set olarak tanımladıkları veri setinde %98,7 doğruluk olarak elde etmişlerdir [186]. Inan vd. veritabanındaki işaret dosyalarının çoğunu kullanmış, veriyi YSA ile test etmiş ve %95,16 sınıflandırma başarımını elde etmiştir [190]. Jenny vd. danışmansız öğrenme yöntemi kullanan bir sınıflandırma algoritması kullandığı için diğer çalışmalardan görece daha başarısız sonuçlar elde etmiştir [191]. Talbi vd. kesirli lineer tahmin yöntemi kullanarak %95 sınıflandırma başarımını elde etmiştir [192].

Bu çalışma herhangi bir öznitelik hesaplamaksızın EKK vurularının sınıflandırılmasında sadece işaret genliğini kullanarak yüksek sınıflandırma başarımını elde edilebileceğini göstermiştir. TBA vuruyu temsil edecek verinin boyutunu azaltmak için

kullanılabilir olduğu görülmüştür. Buna ek olarak TBA öznitelikleri ile elde edilmiş sınıflandırma başarımının GA ile seçilen öznitelikler ile elde edilebildiği görülmüştür. Çalışmada gerçekleştirilen yöntemin diğer algoritmalara göre hızlı çalışmasından dolayı gerçek zamanlı EKK algılaması yapan sistemlere uygulanabileceği düşünülmektedir.

3.2. Dokuz Ayrı EKG Vurusunun Sınıflandırılması ile İlgili Sonuçlar ve Değerlendirme

Çalışmanın ikinci bölümünde EKK vurularının sınıflandırılması için gerçekleştirilen testler daha geniş kapsamda aritmi türlerini kapsayacak şekilde genişletilmiştir. Sekiz aritmi ve bir normal vuru olmak üzere dokuz vuru türünü sınıflandırmak için sınıflandırma algoritmalarının başarımları test edilmiştir. Bir kalp vurusunu temsil edecek öznitelik seti hesaplanmış, bu öznitelik vektörünün boyutunu azaltmak için boyut azaltma yöntemlerinin en iyi sonucu verdiği değerler araştırılmıştır.

Öznitelik vektörlerinin uygun seçimi sınıflama performansını doğrudan etkilemektedir. Vuruların sınıflandırma aşamasında geçilmeden önce öznitelikler belirlenmiş, bir vuruluk işaret 4 ayrı bölgeye ayrılmış ve her bölge için belirlenen istatistiksel öznitelikler hesaplanmıştır. Vurunun tamamı için aynı istatistiksel öznitelikler hesaplanmış ve bu değerlere iki adet R-R aralığı eklenerek öznitelik vektörü oluşturulmuştur.

Hesaplanan öznitelik vektörünün boyutunu azaltmak için TBA, BBA ve ÖH öznitelik azaltma ve SIA, SGA ve GA öznitelik seçim algoritmaları kullanılmıştır. 32 olan öznitelik boyutu bu yöntemler kullanılarak daha düşük boyutlara indirgenmiştir. TBA kullanılarak öznitelik vektörünün TB'leri hesaplanmış ve parametre optimizasyonu yapılarak en iyi başarıyı verecek öznitelik sayısı belirlenmiştir. Hesaplamalar sonucunda en iyi sınıflandırma başarımını veren TB sayısı 17 olarak hesaplanmıştır (Tablo 15).

Benzer şekilde BBA analizi öznitelik vektörünün boyutunu azaltmak için kullanılmış ve BB'ler hesaplanmıştır. En iyi sınıflandırma başarımını verecek BB sayısını bulmak için ızgara arama yöntemi ile belirlenen aralıktaki BB sayıları sınıflandırma işleminde denenmiş ve 9 BB'in en iyi başarımları elde ettiği sonucuna ulaşılmıştır (Tablo 16).

ÖH yöntemi çalışmada kullanılan bir diğer boyut indirgeme yöntemidir. ÖH giriş uzayını daha düşük boyutlu (özellikle 2 veya 3) uzaya düşürmek için kullanılan bir

yöntemdir. Çalışmada ÖH kullanılarak giriş vektör boyutu 2 ve 3 olacak şekilde hesaplamalar yapılmıştır. Ağ boyutu olarak 30-60 aralığında değerlerde testler yapılmış ve bu çıkışları üretecek şekilde deneyler gerçekleştirilmiştir. Elde edilen çıkış değerleri bir sınıflandırıcı ile değerlendirilerek ÖH ile boyut indirgenmiş öznitelikler için sınıflandırma başarımları irdelenmiştir. Bu deneyler sonucunda ÖH ağının çıkış sayısı 3 ve ağın boyutu 45 (Şekil 64) olarak hesaplanmıştır.

TBA, BBA ve ÖH belirlenmiş öznitelik vektörlerinden çeşitli matematiksel yöntemler kullanarak daha düşük boyutlu yeni öznitelik vektörleri hesaplamayı sağlıyordu. GA, SİA ve SGA yöntemleri ise orijinal öznitelikler arasından veriyi en iyi temsil edecek özniteliklerin seçilmesini sağlamaktadırlar. Yani girişte bulunan öznitelikler değiştirilmeden aralarında değersiz olanların elenmesi işlemi gerçekleştirilmektedirler. GA 32 öznitelikten oluşan giriş verisinin boyutunu azaltmak için kullanılmıştır. Hesaplama işleminin hızını arttırmak için GA ile seçilecek özniteliklere alt sınır olarak 5 değeri belirlenmiştir. GA 17 adet özniteliği seçmiş (Tablo 17) ve sınıflandırma işlemlerinde bu öznitelikler kullanılmıştır.

SİA ve SGA yöntemleri kaynak tüketici yöntemler olarak bilinirler ve giriş öznitelik uzayı büyük olduğunda bu yöntemlerin kullanılması imkansız hale gelmektedir. Bu nedenle EKK vuruların sınıflandırılmasında giriş boyutu 200 olduğu için bu iki yöntem ile öznitelik azaltma işlemi gerçekleştirilememiştir. Bu aşamada giriş öznitelik boyutu 32 olduğu için bu iki arama yöntemi öznitelik seçimi için kullanılabilmiştir. SİA yöntemi ile 16 adet öznitelik (Tablo 18) seçilmiş ve öznitelik vektörünün boyutu yarıya düşürülmüştür. Benzer şekilde SGA yöntemi ile 24 öznitelik (Tablo 19) hesaplanmış ve öznitelik vektörünün boyutunda kayda değer azalma sağlamıştır.

Elde edilmiş bu 7 grup öznitelik sınıflandırma algoritmaları tarafından sınıflandırılmış ve sonuçları karşılaştırmalı olarak irdelenmiştir. Sınıflandırma aşamasında her bir deney için 10-kez çapraz doğrulama kullanılmış ve çapraz doğrulama sonucunda elde edilmiş ortalama duyarlılık, kesinlik ve doğruluk değerleri (Eşitlik (82)-(84)) sınıflandırma başarımlarını değerlendirmek için hesaplanmıştır.

Deneylerde ilk olarak çok katmanlı YSA mimarisi kullanılmıştır. Bir adet gizli katmana sahip olan YSA'nın giriş katmanı düğüm sayısı giriş vektörü ile aynı sayıda (32 veya daha az), çıkış katmanı düğüm sayısı 9 olarak belirlenmiştir. Gizli katman düğüm sayısını belirlemek için deneylerde değişik değerlerde düğüm sayısı için testler yapılmış

(Tablo 21) ve 20 civarı değerlerde en iyi sonuçları elde ettiği için gizli katman düğüm sayısı 20 olarak belirlenmiştir. Çalışmada kullanılan YSA geri yayılım algoritması ile eğitilmiştir. YSA en iyi sonucu GA ile seçilmiş öznitelikleri kullanarak %98,94 doğruluk değerini elde etmiştir.

K-NN ile sınıflandırma deneylerinde k parametresinin sınıflandırma başarımına etkisi araştırılmıştır (Tablo 22). Bu amaçla belirlenen aralıktaki tüm k değerleri için sınıflandırma sonuçları irdelenmiştir. K-NN sınıflandırıcı SİA, SGA ve GA ile seçilen öznitelikler için en iyi sınıflandırma başarımına ulaşmış ve en yüksek değerini ileri arama özniteliklerini kullanarak %99,37 doğruluk ile elde etmiştir.

DVM testlerde kullanılan bir diğer sınıflandırıcıdır. DVM ile sınıflandırma işlemleri gerçekleştirilirken DVM'ye ait çekirdek fonksiyonu ve diğer parametrelerin sınıflandırma başarımlarına etkisi araştırılmıştır (Tablo 23-24). Bu deneyler sonucunda çekirdek fonksiyonu olarak *polinom*, polinom derecesi 3, C=1 ve Gamma=1 parametre değerleri en iyi sonucu vermiştir. DVM %98,86 doğruluk ile sınıflandırma başarımı elde etmiştir.

Testlerde kullanılan son sınıflandırıcı KA sınıflandırıcıdır. Önceki sınıflandırma yöntemlerine benzer şekilde parametre optimizasyonu yapılmış ve maksimum derinlik ve minimum yaprak sayısı parametreleri belirlenmeye çalışılmıştır (Tablo 25). Izgara arama işleminden sonra maksimum derinlik parametresi 12 ve minimum yaprak sayısı olarak 4 değeri deneylerde en iyi sonucu vermiştir.

Giriş verisi olarak 32 istatistiksel öznitelik kullanıldığında K-NN ve YSA benzer sınıflandırma başarımlarını elde etmiştir (Tablo 26). KA sınıflandırıcı öznitelik vektör boyutunun büyük olmasından dolayı diğer sınıflandırıcılara nazaran daha kötü sınıflandırma başarımı elde etmiştir. TB'leri giriş verisi olarak kullandığında sınıflandırıcılar bir önceki aşamadaki öznitelikler ile çok yakın sonuçlar elde etmiştir (Tablo 27). KA sınıflandırıcı da kısmen başarımını arttırmıştır. Giriş verisi olarak BB'ler kullanıldığında K-NN sınıflandırıcı diğer sınıflandırıcılara göre daha iyi bir başarımlar elde etmiştir (Tablo 28). ÖH ile öznitelik vektör boyutu 3'e düşürülmüştür. Buna rağmen K-NN sınıflandırıcı yüksek başarımlar elde etmiştir (Tablo 29). Ancak diğer sınıflandırıcıların başarımları önceki öznitelik setlerine göre düşmüştür.

Sınıflandırıcılar en iyi sonuçlarını GA, SİA ve SGA algoritmaları ile seçilen öznitelikleri kullanarak elde etmiştir. GA ile seçilmiş öznitelikleri kullanan K-NN sınıflandırıcı %99,30 ile yüksek sınıflandırma başarımı elde etmiştir. Testlerde bu üç yöntem ile seçilmiş öznitelikler birbirine yakın değerler elde etmiştir. KA sınıflandırıcı en başarılı sonucunu SİA sonucu elde edilmiş öznitelikleri kullanarak %95,20 doğruluk ile elde etmiştir. Yine aynı giriş vektörünü kullanan K-NN sınıflandırıcı %99,37 sınıflandırma başarımı elde ederek deneylerdeki en iyi sonucu elde etmiştir.

Tezde gerçekleştirilen tüm sınıflandırma işlemlerinde 10-kez çapraz doğrulama kullanılmıştır. Örnek olması açısından Tablo 33'de her bir doğrulamada elde edilen sonuç gösterilmiştir. Bu sonuçlardan da görüleceği gibi her bir çapraz doğrulamada farklı sınıflandırma başarımı elde edilmektedir. Birçok doğrulama adımında ortalama sonuçtan daha yüksek sınıflandırma başarımı elde edildiği görülmüştür. Bu nedenle elde edilen sonuçların klasik yöntemler ile elde edilecek sonuçlardan daha tutarlı olduğu aşikardır.

Tezde önerilen yaklaşım deneylerde kullanılan veri seti dışında nasıl sonuç vereceğini incelemek için MIT-BIH aritmi veritabanında bulunan çalışmada belirtilmiş 9 sınıfa ait tüm vurular teste tabi tutulmuş ve sınıflandırma başarımı irdelenmiştir (Tablo 34-35). Tüm vurular üzerinde yapılan testte %99,08 doğruluk ile sınıflandırma başarımı elde edilmiştir.

Tez içerisinde incelenen tüm EKG vuruları 2 numaralı standart bipolar derivasyonla alınmış ve farklı derivasyonların öznitelik hesabında kullanılmasına gerek kalmamıştır. Tek bir derivasyon kullanılıyor olması sistemin işlem yükünü ve problemin karmaşıklığını azaltmıştır. Tezde yapılan çalışmalar Intel I5 2,8GHZ işlemcili ve 10GB DDR3 belleği olan Windows 10 64Bit bir sistemde gerçekleştirilmiş ve ölçülen işlem süreleri elde edilmiştir.

Deneylerde kullanılan sınıflandırma algoritmalarının sınıflandırma süreleri ölçülmüş ve YSA'nın diğer sınıflandırıcılara göre daha yavaş kaldığı gözlemlenmiştir (Tablo 36). Öznitelik azaltma işlemlerinin harcadığı süreler hesaplanmış ve TBA ve BBA algoritmaları en hızlı çözümler olmuştur (Tablo 37). Öznitelik seçim algoritmaları değerlendirme fonksiyonu 10-kez çapraz doğrulama yapan bir sınıflandırıcı olduğu için diğer yöntemlere nazaran daha yavaş kalmıştır. Ancak hesaplamaların bir kez yapılacak olması düşünüldüğünde kabul edilebilir süreler olarak değerlendirilebilir.

Tablo 39. Çalışmada önerilen yöntem ve yayınlanmış EKG aritmi sınıflandırıcılarının yazarları tarafından raporlanan performans metrikleri (%)

Araştırmacılar	Kullanılan Sınıflandırıcı	Doğ.	Duy.	Kes.	Sınıf Sayısı
Liu vd. [193]	Vektör kuantalama ve k-means++	94,6	-	-	8
Thomas vd.[194]	YSA	94,64	88,60	96,18	5
Dilmac vd. [195]	Değiştirilmiş yapay arı kolonisi	99,24	-	-	3
Nazarahari vd. [196]	YSA	98,77	-	-	6
Khalaf vd. [197]	Spektral Korelasyon ve DVM	98,6	99,20	99,7	5
Escalona-Morán vd. [198]	Lojistik Regresyon	98,43	-	97,75	5
Oster vd. [199]	Değiştirilmiş Kalman Filtresi	-	98,3 (F1)	-	4
Paul vd. [200]	Olasılıklı YSA	-	88,33	-	3
Alshraideh vd. [201]	KA	98,29	-	-	7
Sumathi vd. [203]	ANFIS	-	-	99,52	6
Ceylan vd. [204]	Bulanık YSA	99	-	-	10
Melgani vd. [205]	DVM	85,70	-	-	-
Jiang vd. [206]	YSA	98,1	86,6	95,8	5
Güler vd. [207]	Birleştirilmiş YSA	96,94	-	97,78	4
Engin vd. [208]	Bulanık YSA	98	-	-	4
Osowski vd. [209]	Bulanık Hibrit YSA	96	-	-	6
Tezdeki Yaklaşım	K-NN + SİA	99,37	99,01	98,69	9
	YSA + SGA	98,99	98,56	97,78	
	DVM + 32 Özn.	98,92	98,43	97,58	
	KA + SİA	95,20	86,31	84,31	

Tablo 39’da tezde önerilen EKG aritmilerin sınıflandırılması için önerilen yöntemin elde ettiği başarımla literatürde aynı başlıkta yapılmış çalışmalara ait yazarlar tarafından belirtilmiş sonuçların karşılaştırılması gösterilmektedir. Tabloda ilk olarak güncel (2014-2017) çalışmalara yer verilmiştir [189-199]. Tablonun devam eden alt satırlarında [200-205] ise kısmen daha eski ancak en çok atıf alan çalışmalar bulunmaktadır.

Tabloda yazarlar tarafından belirtilen performans metrikleri gösterilmiştir. Her yayında ağırlıklı olarak doğruluk metriğine yer verilmiştir. [204] numaralı çalışmada yazarlar performans ölçütü olarak F1 skorunu beyan etmiştir. Bu nedenle duyarlılık

sütununda F1 skoru yazılmıştır. Ceylan vd. yaptıkları çalışmada 10 ayrı aritmi türünü sınıflandırdıklarını raporlamışlardır. Ancak yazarlar çalışmalarında 173 adet kalp vurusu kullanmışlardır. 10 ayrı sınıf için bu sayı çok az olarak değerlendirilebilir [204]. Yapılan çalışmaların çoğunda ([194,196,200,203,204,206–209]) yazarlar YSA ve türevlerini kullanmıştır. %99,24 sınıflandırma doğruluğu ile tezde yapılan çalışmaya en yakın sonuç Dilmac vd. tarafından yapılan çalışmada elde edilmiştir [195]. Ancak yazarlar bu başarıyı elde etmek için sadece 3 aritmi türünde sınıflandırma yapmıştır. Bir diğer başarılı çalışmada yazarlar sınıflandırma doğruluğu yerine kesinlik ölçütünü karşılaştırma kriteri olarak kullanmışlar ve %99,52 kesinlik başarılarını 6 aritmi türünü sınıflandırarak elde etmiştir. Çalışmalarında DVM kullanan Melgani vd. karşılaştırılan diğer çalışmalar arasında en düşük sınıflandırma sonucunu elde etmiştir [205].

Tablodan görüleceği gibi önerilen yöntem sınıflandırdığı aritmi sayısı ve elde ettiği sınıflandırma başarımları ile yapılan diğer çalışmalardan daha yüksek sınıflandırma başarımları elde etmiştir. Aritmilerin sınıflandırılması için bir vuruluk işareti temsil edecek hesaplanmış öznitelik vektörü sınıflandırma başarımlarını elde etmek için yeterli olmuştur. Çalışmada kullanılan ileri arama, geri arama ve GA öznitelik seçim algoritmaları aritmi sınıflandırmada öznitelik seçimi için uygun öznitelik seçimi algoritmaları olmuştur. K-NN sınıflandırıcı diğer sınıflandırıcılara göre en iyi başarımları elde etmiş ve hesaplama süreleri tüm öznitelik setleri için en hızlı algoritmalar arasında yer almaktadır.

Bilginin daha az veri ile temsil edilmesi her zaman istenen bir sonuçtur. Çalışmada özellikle farklı öznitelik azaltma yöntemlerinin performansları test edilmiştir. Testlerde kullanılan TBA ve BBA ile öznitelik boyutu azaltılmış ve performans artışı sağlanmıştır. Ancak her yeni veri için tekrar hesaplama adımı gerektirmesinden dolayı öznitelik seçim algoritmaları tercih sebebi olmaktadır. Çünkü öznitelik seçim algoritmaları ile değerli öznitelikler bir kez belirlendiğinde her yeni gelen veride değerli olanları filtrelemek yeterli olacaktır.

4. ÖNERİLER

Bu tezde, erken karıncık kasılması virusunu herhangi bir şekilsel öznitelik hesaplamaksızın sınıflandırmak için bir kalp vuruluk EKG işaretine ait zaman serisi öznitelik vektörü olarak kullanılmış ve önerilen sınıflandırma şeması ile yüksek doğrulukta sınıflandırılmıştır. Ek olarak yaygın kullanılan öznitelik azaltma ve seçim algoritmaları ile elde edilen sonuçlar ile performansı karşılaştırılmıştır. Zaman serileri üzerinden GA ile seçilen öznitelikler ile en yüksek başarımlar elde edilmiştir.

Bununla birlikte, deneyler veritabanından alınan kayıtlar üzerinde gerçekleştirilmiş, gerçek zamanlı kayıtlarda uygulaması üzerine herhangi bir çalışma yapılmamıştır. Yapılan çalışmalar bu alanda yapılmış hemen hemen tüm çalışmalarda olduğu gibi MIT-BIH aritmi veritabanı üzerindeki kayıtlar üzerinden gerçekleştirilmiştir. EKK virüsünün algılanması için geliştirilecek bir uygulamaya uyarlanarak gerçek zamanlı çalışması sağlanabilir.

Tezde sunulan ikinci yaklaşımda 9 ayrı aritmi türüne ait sınıflandırma işlemi gerçekleştirecek yeni bir sınıflandırma şeması ve öznitelik seti önerilmiştir. Bu sınıflandırma işleminin gerçekleştirilebilmesi için bir kalp vuruluk EKG işaretine ait zaman serisinden istatistiksel öznitelikler çıkarılmıştır. Bir kalp vuruluk zaman 4 ayrı bölgeye bölünerek 6 ayrı istatistik özniteliği hesaplanmıştır. Öznitelik indirgeme yöntemleri kullanılarak bu özniteliklerin sayısı azaltılmış ve sonuçlar elde edilmiştir. Çalışmada MIT-BIH aritmi veritabanında bulunan en yaygın 9 kalp virüsü sınıflandırılmıştır. Bu sayı artırılabilir ve diğer aritmi türleri çalışmaya dahil edilebilir. Tez çalışması esnasında literatürde bulunan birçok sınıflandırıcı denenmiş ancak sonuçları kötü çıktığından dolayı karşılaştırma listelerine ve teze eklenmemiştir.

Hazırlanacak bilgisayar destekli tanı sistemlerine önerilen yöntemler uygulanarak aritmi sınıflandırma için kullanılabilir. Mobil teknolojilerin gelişimi ile birlikte akıllı telefon ve tabletler gibi cihazlara geliştirilebilecek sağlık uygulamalarına eklenerek kişisel sağlık ön tanı uygulaması olarak çalışabilir. Önerilen sınıflandırma şemaları ve öznitelik setleri geliştirilebilecek taşınabilir EKG kayıt cihazları için ön tanımlı algılayıcı olarak kullanılacak yazılımlar için bilimsel alt yapı oluşturabilir.

5. KAYNAKÇA

1. Collage, O., Anatomy & Physiology. Houston: OpenStax CNX, 2013.
2. Tosun, N., Erkoç, Y., Buzgan, T., Keskinçılıç, B., Aras, D., Yardım, N., Gögen, S., Sarıoğlu, G. ve Soylu, M., Türkiye Kalp ve Damar Hastalıklarının Önleme ve Kontrol Programı (2010-2014), Anıl Matbaası, Ankara, 2014.
3. Go, A. S., Mozaffarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Blaha, M. J., Dai, S., Ford, E. S., Fox, C. S., Franco, S., Fullerton, H. J., Gillespie, C., Hailpern, S. M., Heit, J. A., Howard, V. J., Huffman, M. D., Judd, S. E., Kissela, B. M., Kittner, S. J., Lackland, D. T., Lichtman, J. H., Lisabeth, L. D., Mackey, R. H., Magid, D. J., Marcus, G. M., Marelli, A., Matchar, D. B., McGuire, D. K., Mohler, E. R., Moy, C. S., Mussolino, M. E., Neumar, R. W., Nichol, G., Pandey, D. K., Paynter, N. P., Reeves, M. J., Sorlie, P. D., Stein, J., Towfighi, A., Turan, T. N., Virani, S. S., Wong, N. D., Woo, D. ve Turner, M. B., Heart disease and stroke statistics-2014 update: a report from the American Heart Association. Circulation, 129, 3 (2014) e28–e292.
4. Dokur, Z. ve Ölmez, T., ECG beat classification by a novel hybrid neural network Comput. Methods Programs Biomed., 66, 2–3 (2001) 167–181.
5. Ekşi, A., Zoghi, M. ve Çertuğ, A., Hastane Öncesi Acil Bakımda EKG ve Aritmi Yönetimi, 4. Baskı. İzmir: Ege Üniversitesi Basımevi, 2014.
6. Rangayyan, R. M., Biomedical Signal Analysis: A Case-Study Approach. IEEE Press, 2002.
7. Einthoven, W., Ueber die Form des menschlichen Electrocardiogramms Pflügers Arch. f. Physiol., 60 (1885) 101–123.
8. Electrocardiography <https://en.wikipedia.org/wiki/Electrocardiography> 1 Ocak 2016.
9. ECG library <http://www.ecglibrary.com/ecghist.html> 1 Ocak 2016.
10. Barnes, A., Standardization of precordial leads Supplementary report Am. Heart J., 15, 2 (1938) 235–239.
11. Bruce, R. (University of W. D. of A. C., Blackmon, J. (University of W. D. of A. C., Jones, J. (University of W. D. of A. C. ve Strait, G. (University of W. D. of A. C., Exercise Testing in Adult Normal Subjects and Cardiac Patients Pediatrics, 32 (1963) 742–756.
12. Thakor, N. V, Webster, J. G. ve Tompkins, W. J., Estimation of QRS complex power spectra for design of a QRS filter. IEEE Trans. Biomed. Eng., 31 (1984) 702–706.
13. Clifford, G., Azuaje, F. ve McSharry, P., Advanced methods and tools for ECG data analysis. 2006.

14. Moody, G. ve Mark, R., The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it, [1990] Proceedings Computers in Cardiology, 1990, IEEE Comput. Soc. Press, 185–188.
15. Gholam-Hosseini, H., Nazeran, H. ve Reynolds, K. J., ECG noise cancellation using digital filters, Proceedings of the 2nd International Conference on Bioelectromagnetism (Cat. No.98TH8269), 1998, IEEE, 151–152.
16. Łęski, J. M. ve Henzel, N., ECG baseline wander and powerline interference reduction using nonlinear filter bank Signal Processing, 85 (2005) 781–793.
17. Jane, R., Laguna, P., Thakor, N. V. ve Caminal, P., Adaptive baseline wander removal in the ECG: Comparative analysis with cubic spline technique, Proceedings Computers in Cardiology, 1992, IEEE Comput. Soc. Press, 143–146.
18. Mneimneh, M. A., Yaz, E. E., Johnson, M. T. ve Povinelli, R. J., An adaptive kalman filter for removing baseline wandering in ECG signals 2006 Comput. Cardiol., (2006).
19. Sornmo, L., Time-varying filtering for removal of baseline wander in exercise ECGs [1991] Proc. Comput. Cardiol., (1991).
20. Barati, Z. ve Ayatollahi, A., Baseline Wandering Removal by Using Independent Component Analysis to Single-Channel ECG data 2006 Int. Conf. Biomed. Pharm. Eng., (2006).
21. Boucheham, B., Ferdi, Y. ve Batouche, M. C., Piecewise linear correction of ECG baseline wander: A curve simplification approach Comput. Methods Programs Biomed., 78 (2005) 1–10.
22. Zhang, D., Wavelet Approach for ECG Baseline Wander Correction and Noise Reduction. Conf. Proc. IEEE Eng. Med. Biol. Soc., 2 (2005) 1212–1215.
23. Friesen, G. M., Jannett, T. C., Jadallah, M. a, Yates, S. L., Quint, S. R. ve Nagle, H. T., A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Trans. Biomed. Eng., 37, 1 (1990) 85–98.
24. Kaya, Y. ve Pehlivan, H., Feature Selection Using Genetic Algorithms for Premature Ventricular Contraction Classification, 9th International Conference on Electrical and Electronics Engineering, ELECO 2015, 2015.
25. Kaya, Y. ve Pehlivan, H., Comparison of classification algorithms in classification of ECG beats by time series, 2015 23rd Signal Processing and Communications Applications Conference (SIU), 2015, IEEE, 407–410.
26. Kaya, Y. ve Pehlivan, H., KNN, NN, Bayes, DT ve SVM Kullanılarak EKG Vurularının Sınıflandırılması, 2ND International Symposium on Innovative Technologies In Engineering and Science, 2014, 1–10.
27. Kaya, Y., Pehlivan, H. ve Tenekeci, M. E., Effective ECG beat classification using higher order statistic features and genetic feature selection Biomed. Res., 28, 17 (2017) 7594–7603.

28. Saxena, S. C., Sharma, A. ve Chaudhary, S. C., Data compression and feature extraction of ECG signals Int. J. Syst. Sci., 28, 5 (1997) 483–498.
29. Elghazzawi, Z. ve Geheb, F., A knowledge-based system for arrhythmia detection, Computers in Cardiology 1996, 1996, IEEE, 541–544.
30. Liang, H., ECG Feature Elements Identification For Cardiologist Expert Diagnosis. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., 4 (2005) 3845–8.
31. Ahlstrom, M. L. ve Tompkins, W. J., Automated High-Speed Analysis of Holter Tapes with Microcomputers IEEE Trans. Biomed. Eng., BME-30, 10 (1983) 651–657.
32. Abnormalities in the ECG Measurements <http://ecg.utah.edu/lesson/4> .
33. Gholam-Hosseini, H. ve Nazeran, H., Detection and extraction of the ECG signal parameters, Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286), 1998, IEEE, 127–130.
34. Segura-Juárez, J. J., Cuesta-Frau, D. ve Samblas-Pena, L., A microcontroller-based portable electrocardiograph system, Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems, 2003, II: 922–925.
35. Jose Luis Rodriguez Sotelo, Biosignal analysis for cardiac arrhythmia detection using non-supervised techniques, Universidad Nacional de Colombia, 2010.
36. de Chazal, P., O'Dwyer, M. ve Reilly, R. B., Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Trans. Biomed. Eng., 51, 7 (2004) 1196–206.
37. Moody, G. B. ve Mark, R. G., The impact of the MIT-BIH arrhythmia database. IEEE Eng. Med. Biol. Mag., 20, 3 (2001) 45–50.
38. Kutlu, Y., Multi-Stage Classification of Abnormal Patterns in EEG and ECG using Model-Free Methods, Dokuz Eylül University, 2010.
39. Crawford, M. H. ve Dimarco, J. P., Crawford Kardiyoloji, 1. Baskı. İstanbul: AND Danışmanlık, Eğitim, Yayıncılık, 2004.
40. Roberts, D. M. J., Signals and Systems: Analysis Using Transform Methods and MATLAB. 2004.
41. Prasad, G. K. ve Sahambi, J. S., Classification of ECG arrhythmias using multi-resolution analysis and neural networks, TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region, 2003, Allied Publishers Pvt. Ltd I: 227–231.
42. Ahlstrom, M. L. ve Tompkins, W. J., Digital filters for real-time ECG signal processing using microprocessors. IEEE Trans. Biomed. Eng., 32, 9 (1985) 708–13.

43. Hamilton, P. S., 60Hz filtering for ECG signals: to adapt or not to adapt?, Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1993, IEEE, 779–780.
44. Jane, R., Laguna, P., Caminal, P. ve Rix, H., Adaptive filtering of high-resolution ECG signals [1990] Proc. Comput. Cardiol., (1990).
45. Kumari, R. S. S., Bharathi, S. ve Sadasivam, V., Design of Optimal Discrete Wavelet for ECG Signal Using Orthogonal Filter Bank Int. Conf. Comput. Intell. Multimed. Appl. (ICCIMA 2007), 1 (2007).
46. Kuzume, K. ve Niiijima, K., Optimization of lifting wavelet filters for ECG data compression ICSP '98. 1998 Fourth Int. Conf. Signal Process. (Cat. No.98TH8344), (1998).
47. Laguna, P., Jane, R. ve Caminal, P., Adaptive filtering of ECG baseline wander 1992 14th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2 (1992) 508–509.
48. Nikolaev, N., Nikolov, Z., Gotchev, A. ve Egiazarian, K., Wavelet domain Wiener filtering for ECG denoising using improved signal estimate 2000 IEEE Int. Conf. Acoust. Speech, Signal Process. Proc. (Cat. No.00CH37100), 6 (2000).
49. Oltean, M., Boucher, J.-M. ve Isar, A., Map Filtering in the Diversity-Enhanced Wavelet Domain Applied to Ecg Signals Denoising 2006 IEEE Int. Conf. Acoust. Speech Signal Process. Proc., 2 (2006).
50. Sameni, R., Shamsollahi, M. B., Jutten, C. ve Babaie-Zade, M., Filtering noisy ECG signals using the extended kalman filter based on a modified dynamic ECG model, Computers in Cardiology, 2005, IEEE, 1017–1020.
51. Sameni, R., Shamsollahi, M. B., Jutten, C. ve Clifford, G. D., A nonlinear Bayesian filtering framework for ECG denoising IEEE Trans. Biomed. Eng., 54 (2007) 2172–2185.
52. Shi, Z. S. Z., Bao, Z. B. Z., Jiao, L. J. L. ve Ma, K. M. K., Nonlinear ECG filtering by group normalized wavelet packets Proc. IEEE Int. Symp. Inf. Theory, (1997).
53. Ge, J., Chen, H. ve Xu, Z., Study of high-frequency electrocardiogram, Annual Reports of the Research Reactor Institute, Kyoto University, 2001, I: 32–35.
54. Pettersson, J., Carro, E., Edenbrandt, L., Pahlm, O., Sornmo, L., Warren, S. G. ve Wagner, G. S., High frequency ECG during acute myocardial ischemia Comput. Cardiol. 1997, (1997).
55. Abboud, S., High frequency ECG-a new method to examine depolarization changes mediated by transient myocardial ischemia [1989] Proceedings. Comput. Cardiol., (1989).
56. Ewing, D. J., Neilson, J. M. ve Travis, P., New method for assessing cardiac parasympathetic activity using 24 hour electrocardiograms. Br. Heart J., 52 (1984) 396–402.

57. Petelenz, T., Flak, Z., Czyz, Z., Sosnowski, M., Slominska-Petelenz, T. ve Leski, J., High frequency electrocardiogram in detection of previous myocardial infarction: a noninvasive test complementary to routine electrocardiogram Proc. Comput. Cardiol. Conf., (1993).
58. Addison, P. S., Watson, J. N., Clegg, G. R., Holzer, M., Sterz, F. ve Robertson, C. E., Evaluating arrhythmias in ECG signals using wavelet transforms IEEE Eng. Med. Biol. Mag., 19, 5 (2000) 104–109.
59. Castro, B., Kogan, D. ve Geva, A. B., ECG feature extraction using optimal mother wavelet 21st IEEE Conv. Electr. Electron. Eng. Isr. Proc. (Cat. No.00EX377), (2000).
60. Mahmoodabadi, S., Ahmadian, A., Abolhasani, M., Eslami, M. ve Bidgoli, J., ECG Feature Extraction Based on Multiresolution Wavelet Transform. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., 4 (2005) 3902–5.
61. Matsuyama, A. ve Jonkman, M., The application of wavelet and feature vectors to ECG signals. Australas. Phys. Eng. Sci. Med., 29, 1 (2006) 13–7.
62. Anant, K., Dowla, F. ve Rodrigue, G., Vector quantization of ECG wavelet coefficients IEEE Signal Process. Lett., 2, 7 (1995) 129–131.
63. Brouse, C., Dumont, G. A., Herrmann, F. J. ve Ansermino, J. M., A wavelet approach to detecting electrocautery noise in the ECG IEEE Eng. Med. Biol. Mag., 25 (2006) 76–82.
64. Ye Datian ve Ouyang Xuemei, Application of wavelet analysis in detection of fetal ECG, Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1996, IEEE III: 1043–1044.
65. Gamo, C., An implementation of the wavelet transform for ECG analysis, First International Conference on Advances in Medical Signal and Information Processing, 2000, IEE MM: 32–40.
66. Li, B.-L. L. B.-L. ve Wu, H. W. H., Wavelet transformation of chaotic biological signals, Proceedings of the 1995 Fourteenth Southern Biomedical Engineering Conference, 1995, IEEE, 185–188.
67. Cuiwei Li, Chongxun Zheng ve Changfeng Tai, Detection of ECG characteristic points using wavelet transforms IEEE Trans. Biomed. Eng., 42, 1 (1995) 21–28.
68. Martínez, J. P., Almeida, R., Olmos, S., Rocha, A. P. ve Laguna, P., A wavelet-based ECG delineator: evaluation on standard databases. IEEE Trans. Biomed. Eng., 51, 4 (2004) 570–81.
69. Misiti, M. ve Misiti, Y., Wavelet toolbox The MathWorks Inc., (1996).
70. Walnut, D. F., An Introduction to Wavelet Analysis, 2nd ed. Basel: Springer, 2003.

71. H Karel, J., M Peeters, R., Westra, R., S Moermans, K., P Haddad, S. ve Serdijn, W., Optimal discrete wavelet design for cardiac signal processing. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., 3 (2005) 2769–72.
72. Szildgyi, S. M., Szildgyi, L. ve David, L., Comparison Between Neural-Network-Based Adaptive Filtering and Wavelet Transform for Ecg Characteristic Points Detection, Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. “Magnificent Milestones and Emerging Opportunities in Medical Engineering” (Cat. No.97CH36136), 1997, IEEE I: 272–274.
73. Popescu, M., Cristea, P. ve Bezerianos, A., High resolution ECG filtering using adaptive Bayesian wavelet shrinkage, Computers in Cardiology 1998. Vol. 25 (Cat. No.98CH36292), 1998, IEEE, 401–404.
74. Bezerianos, A., Popescu, M., Laskaris, N., Manolis, A., Hiladakis, I., Stathopoulos, C. ve Cristea, P., Selective noise filtering of high resolution ECG through wavelettransform Comput. Cardiol. 1996, (1996).
75. Provaznik, I. ve Kozumplik, J., Wavelet-based compression of ECG signals Proc. 18th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 3 (1996).
76. Moghaddam, A. R. A. ve Nayebi, K., A two dimensional wavelet packet approach for ECG compression, Proceedings of the Sixth International Symposium on Signal Processing and its Applications (Cat.No.01EX467), 2001, IEEE I: 226–229.
77. Sastry, R. V. S. ve Rajgopal, K., ECG compression using wavelet transform, Proceedings of the First Regional Conference, IEEE Engineering in Medicine and Biology Society and 14th Conference of the Biomedical Engineering Society of India. An International Meet, 1995, IEEE, 2/52-2/53.
78. Bilgin, A., Marcellin, M. W. ve Altbach, M. I., Wavelet compression of ECG signals by JPEG2000, Data Compression Conference, 2004. Proceedings. DCC 2004, 2004, IEEE, 527–527.
79. Miaou, S.-G. ve Chao, S.-N., Wavelet-based lossy-to-lossless ECG compression in a unified vector quantization framework. IEEE Trans. Biomed. Eng., 52, 3 (2005) 539–43.
80. Tuzman, A. ve Chialanza, S., Design of wavelet basis for ECG compression, Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286), 1998, IEEE, 198–201.
81. Hao, Y. H. Y. ve Marziliano, P., An efficient wavelet-based pattern matching scheme for ECG data compression IEEE Int. Work. Biomed. Circuits Syst. 2004., (2004).
82. Alesanco, A., Olmos, S., Istepanian, R. ve Garcia, J., A novel real-time multilead ECG compression and de-noising method based on the wavelet transform, Computers in Cardiology, 2003, 2003, IEEE, 593–596.

83. Yen, H.-L. Y. H.-L. ve Miaou, S.-G. M. S.-G., ECG compression using dynamic tree vector quantization in wavelet domain 2001 Conf. Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2 (2001).
84. Qibin Zhao ve Liqing Zhang, ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines, 2005 International Conference on Neural Networks and Brain, 2005, IEEE II: 1089–1092.
85. Matsuyama, A., Jonkman, M. ve de Boer, F., Improved ECG signal analysis using wavelet and feature extraction. Methods Inf. Med., 46, 2 (2007) 227–30.
86. Börjesson, P. O., Pahlm, O., Sörnmo, L. ve Nygård, M. E., Adaptive QRS detection based on maximum a posteriori estimation. IEEE Trans. Biomed. Eng., 29, 5 (1982) 341–51.
87. Kohler, B.-U., Hennig, C. ve Orglmeister, R., The principles of software QRS detection IEEE Eng. Med. Biol. Mag., 21, 1 (2002) 42–57.
88. Taddei, A., Biagini, A., Distanti, G., Emdin, M., Mazzei, M. G., Pisani, P., Roggero, N., Varanini, M., Mark, R. G., Moody, G. B., Braaksma, L., Zeelenberg, C. ve Marchesi, C., The European ST-T database: development, distribution and use, [1990] Proceedings Computers in Cardiology, 1990, IEEE Comput. Soc. Press, 177–180.
89. Lepage, R., Boucher, J.-M., Blanc, J.-J. ve Cornilly, J.-C., ECG segmentation and P-wave feature extraction: application to patients prone to atrial fibrillation, 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2001, IEEE I: 298–301.
90. Syeda-Mahmood, T., Beymer, D. ve Wang, F., Shape-based matching of ECG recordings Conf. Proc. IEEE Eng. Med. Biol. Soc., 2007 (2007) 2012–8.
91. Szilagy, S. M., Benyo, Z., Szilagy, L. ve David, L., Adaptive wavelet-transform-based ECG waveforms detection, Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439), 2003, IEEE III: 2412–2415.
92. Wong, S., Ng, F., Mora, F., Passariello, G. ve Almeida, D., QT interval time frequency analysis using Haar wavelet, Computers in Cardiology 1998. Vol. 25 (Cat. No.98CH36292), 1998, IEEE, 405–408.
93. Gutierrez, A., Hernandez, P., Lara, M. ve Perez, S., A QRS detection algorithm based on Haar wavelet, Computers in Cardiology 1998. 25 (Cat. No.98CH36292), 1998, IEEE, 353–356.
94. Fraden, J. ve Neuman, M. R., QRS wave detection Med. Biol. Eng. Comput., 18, 2 (1980) 125–132.
95. Ligtenberg, A. ve Kunt, M., A robust-digital QRS-detection algorithm for arrhythmia monitoring Comput. Biomed. Res., 16, 3 (1983) 273–286.

96. Goovaerts, H. G., Ros, H. H., Van Den Akker, T. J. ve Schneider, H., A Digital QRS Detector Based on the Principle of Contour Limiting IEEE Trans. Biomed. Eng., BME-23, 2 (1976) 154–160.
97. Yanowitz, F., Kinias, P., Rrawling, D. ve Fozzard, H. A., Accuracy of a Continuous Real-Time ECG Dysrhythmia Monitoring System Circulation, 50, 1 (1974) 65–72.
98. Cox, J. R., Nolle, F. M., Fozzard, H. A. ve Oliver, G. C., AZTEC, a Preprocessing Program for Real-Time ECG Rhythm Analysis IEEE Trans. Biomed. Eng., BME-15, 2 (1968) 128–129.
99. Kadambe, S., Murray, R. ve Boudreaux-Bartels, G. F., Wavelet transform-based QRS complex detector IEEE Trans. Biomed. Eng., 46, 7 (1999) 838–848.
100. Dinh, H. A. N., Kumar, D. K., Pah, N. D. ve Burton, P., Wavelets for QRS detection, 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2001, IEEE II: 1883–1887.
101. Ghaffari, A., Golbayani, H. ve Ghasemi, M., A new mathematical based QRS detector using continuous wavelet transform Comput. Electr. Eng., 34, 2 (2008) 81–91.
102. Chen, S.-W., Chen, H.-C. ve Chan, H.-L., A real-time QRS detection method based on moving-averaging incorporating with wavelet denoising. Comput. Methods Programs Biomed., 82, 3 (2006) 187–95.
103. Szilagyi, L., Benyo, Z., Szilagyi, S. M., Szlavecz, A. ve Nagy, L., On-line QRS complex detection using wavelet filtering, 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2001, IEEE II: 1872–1874.
104. Chouakri, S. A., Bereksi-Reguig, F. ve Taleb-Ahmed, A., QRS complex detection based on multi wavelet packet decomposition Appl. Math. Comput., 217, 23 (2011) 9508–9525.
105. Xu, X. ve Liu, Y., Adaptive Threshold for QRS Complex Detection Based on Wavelet Transform. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., 7 (2005) 7281–4.
106. Fard, P. J. M., Moradi, M. H. ve Tajvidi, M. R., A novel approach in R peak detection using Hybrid Complex Wavelet (HCW). Int. J. Cardiol., 124, 2 (2008) 250–3.
107. Xue, Q., Hu, Y. H. ve Tompkins, W. J., Neural-network-based adaptive matched filtering for QRS detection IEEE Trans. Biomed. Eng., 39 (1992) 317–329.
108. Suzuki, Y., Self-organizing QRS-wave recognition in ECG using neural networks IEEE Trans. Neural Networks, 6 (1995) 1469–1477.
109. Reaz, M. B. I. ve Wei, L. S., Detection of the R wave peak of QRS complex using neural network Proceedings. 2004 Int. Conf. Inf. Commun. Technol. From Theory to Appl. 2004., (2004).

110. Hamilton, P. S. ve Tompkins, W. J., Adaptive matched filtering for QRS detection Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., (1988).
111. Trahanias, P. E., An approach to QRS complex detection using mathematical morphology. IEEE Trans. Biomed. Eng., 40, 2 (1993) 201–5.
112. Ebenezer, D. ve Krishnamurthy, V., Wave digital matched filter for electrocardiogram preprocessing J. Biomed. Eng., 15, 2 (1993) 132–134.
113. Moody, G. B. ve Mark, R. G., QRS morphology representation and noise estimation using the Karhunen-Loeve transform, [1989] Proceedings. Computers in Cardiology, 1989, IEEE Comput. Soc. Press, 269–272.
114. Kaplan, D. T., Simultaneous QRS detection and feature extraction using simple matched filter basis functions, [1990] Proceedings Computers in Cardiology, 1990, IEEE Comput. Soc. Press, 503–506.
115. Engelse, W. A. H. ve Zeelenberg, C., A single scan algorithm for QRS-detection and feature extraction Comput. Cardiol., 6 (1979) 37–42.
116. Hamilton, P. S. ve Tompkins, W. J., Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans. Biomed. Eng., 33 (1986) 1157–1165.
117. Holsinger, W. P., Kempner, K. M. ve Miller, M. H., A QRS Preprocessor Based on Digital Differentiation IEEE Trans. Biomed. Eng., BME-18 (1971).
118. Lee, R.-G., Chou, I.-C., Lai, C.-C., Liu, M.-H. ve Chiu, M.-J., A Novel QRS Detection Algorithm Applied to The Analysis for Heart Rate Variability of Patients with Sleep Apnea Biomed. Eng. Appl. Basis Commun., 17, 5 (2005) 258–262.
119. Pan, J. ve Tompkins, W. J., A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng., 32, 3 (1985) 230–6.
120. So, H. H. ve Chan, K. L., Development of QRS detection method for real-time ambulatory cardiac monitor Proc. 19th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 'Magnificent Milestones Emerg. Oppor. Med. Eng. (Cat. No.97CH36136), 1 (1997).
121. Wrublewski, T. A., Sun, Y. ve Beyer, J. A., Real-time early detection of R waves of ECG signals Images Twenty-First Century. Proc. Annu. Int. Eng. Med. Biol. Soc., (1989).
122. Zidelmal, Z., Amirou, A., Adnane, M. ve Belouchrani, A., QRS detection based on wavelet coefficients. Comput. Methods Programs Biomed., 107, 3 (2012) 490–6.
123. Dohare, A. K., Kumar, V. ve Kumar, R., An efficient new method for the detection of QRS in electrocardiogram Comput. Electr. Eng., 40, 5 (2014) 1717–1730.
124. Boutana, D., Benidir, M. ve Bouaziz, F., Multiresolution wavelet-based QRS complex detection algorithm suited to several abnormal morphologies IET Signal Process., 8, 7 (2014) 774–782.

125. Provaznik, I., Kozumplik, J., Bardonova, J., Novakova, M. ve Novakova, Z., Wavelet Transform In ECG Signal Processing, EuroConference BIOSIGNAL 2000, 2000, 1–5.
126. Torkkola, K., Feature Extraction by Non-Parametric Mutual Information Maximization J. Mach. Learn. Res., 3 (2003) 1415–1438.
127. Chow, T. ve Huang, D., Estimating Optimal Feature Subsets Using Efficient Estimation of High-Dimensional Mutual Information IEEE Trans. Neural Networks, 16, 1 (2005) 213–224.
128. Tan, K., Chan, K. ve Choi, K., Detection of the QRS complex, P wave and T wave in Electrocardiogram, First International Conference on Advances in Medical Signal and Information Processing, 2000, 41–47.
129. Sun, Y., Chan, K. L. ve Krishnan, S. M., Characteristic wave detection in ECG signal using morphological transform. BMC Cardiovasc. Disord., 5 (2005) 28.
130. Chouhan, V. ve Mehta, S., Total removal of baseline drift from ECG signal Comput. Theory Appl., (2007).
131. Senhadji, L., Wang, F., Hernandez, A. I. ve Carrault, G., Wavelets extrema representation for QRS-T cancellation and P wave detection Comput. Cardiol., (2002).
132. Reddy, B. R. S., Elko, P. E., Christenson, D. W. ve Rowlandson, G. I., Detection of P waves in resting ECG: a preliminary study Proc. Comput. Cardiol., (1992).
133. Freeman, K. ve Singh, A., P Wave Detection Of Ambulatory ECG Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 13, (1991).
134. Yeh, Y. C. ve Wang, W. J., QRS complexes detection for ECG signal: The Difference Operation Method Comput. Methods Programs Biomed., 91 (2008) 245–254.
135. Nazeran, H., Chatlapalli, S. ve Krishnam, R., Effect of Novel Nanoscale Energy Patches on Spectral and Nonlinear Dynamic Features of Heart Rate Variability Signals in Healthy Individuals during Rest and Exercise. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., 5 (2005) 5563–7.
136. Bahoura, M., Hassani, M. ve Hubin, M., DSP implementation of wavelet transform for real time ECG wave forms detection and heart rate analysis Comput. Methods Programs Biomed., 52 (1997) 35–44.
137. Matsuyama, A., ECG and APG signal analysis during exercise in a hot environment, Charles Darwin University, 2009.
138. Forester, J., Bo, H., Sleigh, J. W. ve Henderson, J. D., Variability of R-R, P wave-to-R wave, and R wave-to-T wave intervals. Am. J. Physiol., 273 (1997) H2857–H2860.
139. Karris, S. T., Signals and Systems with MATLAB® Applications. 2003.

140. Chen, G. ve Friedman, E., An RLC interconnect model based on fourier analysis Comput. Des. Integr. Circuits Syst. IEEE Trans., 24 (2005) 170–183.
141. Lin, C. H., Du, Y. C., Chen, Y. F. ve Chen, T. S., Multiple ECG beats recognition in the frequency domain using grey relational analysis Conf Proc IEEE Eng Med Biol Soc, 1 (2006) 2154–2158.
142. Sahambi, J. S., Tandon, S. N. ve Bhatt, R. K. P., Using wavelet transforms for ECG characterization. An on-line digital signal processing system IEEE Eng. Med. Biol. Mag., 16 (1997).
143. Xi, Q., Sahakian, A. V., Ng, J. ve Swiryn, S., Stationarity of surface ECG atrial fibrillatory wave characteristics in the time and frequency domains in clinically stable patients Comput. Cardiol. 2003, (2003).
144. Jenkins, J., Noh, K. H., Guezennec, A., Bump, T. ve Arzbaecher, R., Diagnosis of atrial fibrillation using electrograms from chronic leads: evaluation of computer algorithms. Pacing Clin. Electrophysiol., 11 (1988) 622–631.
145. Dzwonczyk, R., Neuman, R. ve Brown, C., Frequency analysis of the human ECG during ventricular fibrillation Proc. Fifteenth Annu. Northeast Bioeng. Conf., (1989).
146. Nowak, C. N., Fischer, G., Wieser, L., Tilg, B. ve Strohmeier, H. U., Frequency spectrum of the intracardiac and body surface ECG during ventricular fibrillation - a computer model study 2006 Comput. Cardiol., (2006).
147. Afonso, V. X. ve Tompkins, W. J., Detecting ventricular fibrillation IEEE Eng. Med. Biol. Mag., 14 (1995).
148. Abboud, S., Cohen, R. J., Selwyn, A., Ganz, P., Sadeh, D. ve Friedman, P. L., Detection of transient myocardial ischemia by computer analysis of standard and signal-averaged high-frequency electrocardiograms in patients undergoing percutaneous transluminal coronary angioplasty. Circulation, 76 (1987) 585–596.
149. Burrus, C. S., Gopinath, R. A. ve Guo, H., Introduction to Wavelets and Wavelet Transforms: A Primer. New Jersey: Prentice Hall, 1998.
150. Unser, M. ve Aldroubi, A., A review of wavelets in biomedical applications Proc. IEEE, 84 (1996) 626–638.
151. Addison, P. S., The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance : : Amazon.com: Books, 1st ed. CRC Press, 2002.
152. Keselbrener, L., Keselbrener, M. ve Akselrod, S., Nonlinear high pass filter for R-wave detection in ECG signal. Med. Eng. Phys., 19, 5 (1997) 481–4.
153. Tinati, M. A. ve Mozaffary, B., A wavelet packets approach to electrocardiograph baseline drift cancellation. Int. J. Biomed. Imaging, 2006 (2006) 97157.

154. Kudo, M. ve Sklansky, J., Comparison of algorithms that select features for pattern classifiers Pattern Recognit., 33, 1 (2000) 25–41.
155. Yıldız, A., EEG ve EKG İşaretlerinden Örüntü Tanıma Uygulamaları ve Karşılaştırılması, Fırat Üniversitesi, 2011.
156. Somol, P., Pudil, P. ve Kittler, J., Fast branch & bound algorithms for optimal feature selection IEEE Trans. Pattern Anal. Mach. Intell., 26, 7 (2004) 900–912.
157. Somol, P., Novovicova, J. ve Pudil, P., Efficient Feature Subset Selection and Subset Size Optimization, in Pattern Recognition Recent Advances, InTech, 2010 1–24.
158. Jain, A. K., Duin, P. W. ve Jianchang Mao, Statistical pattern recognition: a review IEEE Trans. Pattern Anal. Mach. Intell., 22, 1 (2000) 4–37.
159. Holland, J. H., Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, 1, 1. 1975.
160. Raymer, M. L., Punch, W. F., Goodman, E. D., Kuhn, L. A. ve Jain, A. K., Dimensionality reduction using genetic algorithms IEEE Trans. Evol. Comput., 4, 2 (2000) 164–171.
161. Fukunaga, K., Introduction to Statistical Pattern Recognition, 22, 1990.
162. Everson, R. ve Roberts, S., Independent component analysis: A flexible nonlinearity and decorrelating manifold approach. Neural Comput., 11, 8 (1999) 1957–83.
163. Kantardzic, M., Data Mining: Concepts, Models, Methods, and Algorithms, 2nd Editio. John Wiley & Sons, Inc., 2011.
164. Theodoridis, S. ve Koutroumbas, K., Pattern Recognition, 4. Edition. Burlington: Academic Press, 2009.
165. Hyvärinen, a ve Oja, E., Independent component analysis: algorithms and applications. Neural Netw., 13 (2000) 411–30.
166. Martis, R. J., Acharya, U. R. ve Min, L. C., ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform Biomed. Signal Process. Control, 8, 5 (2013) 437–448.
167. Hyvärinen, A. ve Oja, E., A Fast Fixed-Point Algorithm for Independent Component Analysis Neural Comput., 9, 7 (1997) 1483–1492.
168. Kohonen, T., The self-organizing map Proc. IEEE, 78, 9 (1990) 1464–1480.
169. Kohonen, T., Self-organized formation of topologically correct feature maps Biol. Cybern., 43, 1 (1982) 59–69.
170. Duda, R. O., Hart, P. E. ve Stork, D. G., Pattern Classification, Second Edi. New York: JOHN WILEY & SONS, INC., 2000.

171. Cover, T. ve Hart, P., Nearest neighbor pattern classification IEEE Trans. Inf. Theory, 13, 1 (1967) 21–27.
172. Fausett, L. V., Fundamentals of Neural Networks: Architectures, Algorithms And Applications, 1. edition. Pearson, 1993.
173. Haykin, S., Neural networks: a comprehensive foundation, Second Edi. Ontario: Pearson, 1999.
174. Dokur, Z., Yapay Sinir Ağları ve Genetik Algoritmalar Kullanılarak EKG Vurularının Sınıflandırılması, İstanbul Teknik Üniversitesi, 1999.
175. Bishop, C. M., Neural Networks for Pattern Recognition, 92. Birmingham: CLARENDONPRESS OXFORD, 1995.
176. Lippmann, R., An introduction to computing with neural nets IEEE ASSP Mag., 4, 2 (1987) 4–22.
177. Bishop, C. M., Pattern Recognition and Machine Learning, 4, 4. Springer, 2006.
178. Çomak, E., Destek Vektör Makinelerinin Eğitimi İçin Yeni Yaklaşımlar, Selçuk Üniversitesi, 2008.
179. Vapnik, V. ve Lerner, a, Pattern recognition using generalized portrait method Autom. Remote Control, 24 (1963) 774–780.
180. Cortes, C. ve Vapnik, V., Support-vector networks Mach. Learn., 20, 3 (1995) 273–297.
181. Cherkassky, V. ve Mulier, F. M., Learning from Data: Concepts, Theory, and Methods. Wiley, 2007.
182. Rokach, L. ve Maimon, O., Data mining with decision trees: theory and applications. Singapore: World Scientific Publishing Co., 2008.
183. Han, J., Kamber, M. ve Pei, J., Data Mining: Concepts and Techniques. 2012.
184. De Oliveira, L. S. C., Andreão, R. V ve Sarcinelli-Filho, M., Premature Ventricular beat classification using a dynamic Bayesian Network. Conf. Proc. IEEE Eng. Med. Biol. Soc., 2011 (2011) 4984–7.
185. Ittaturut, S., Lek-Uthai, A. ve Teeramongkonrasmee, A., Detection of Premature Ventricular Contraction for real-time applications, 2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2013, 2013.
186. Bortolan, G., Jekova, I. ve Christov, I., Comparison of four methods for premature ventricular contraction and normal beat clustering, Computers in Cardiology, 2005, XXXII: 921–924.

187. Li, P., Liu, C., Wang, X., Zheng, D., Li, Y. ve Liu, C., A low-complexity data-adaptive approach for premature ventricular contraction recognition Signal, Image Video Process., 8, 1 (2013) 111–120.
188. Ribeiro, B. R., Henirques, J. H., Marques, A. M. ve Antunes, M. A., Manifold learning for premature ventricular contraction detection, *Computers in Cardiology*, 2008, XXXV: 917–920.
189. Foo, S. Y., Stuart, G., Harvey, B. ve Meyer-Baese, A., Neural network-based EKG pattern recognition Eng. Appl. Artif. Intell., 15, 3–4 (2002) 253–260.
190. Inan, O. T., Giovangrandi, L. ve Kovacs, G. T. A., Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. IEEE Trans. Biomed. Eng., 53, 12 Pt 1 (2006) 2507–15.
191. Jenny, N. Z. N., Faust, O. ve Yu, W., Automated Classification of Normal and Premature Ventricular Contractions in Electrocardiogram Signals J. Med. Imaging Heal. Informatics, 4, 6 (2014) 886–892.
192. Talbi, M. L. ve Ravier, P., Detection of PVC in ECG signals using fractional linear prediction Biomed. Signal Process. Control, 23 (2016) 42–51.
193. Liu, T., Si, Y., Wen, D., Zang, M. ve Lang, L., Dictionary learning for VQ feature extraction in ECG beats classification Expert Syst. Appl., 53 (2016) 129–137.
194. Thomas, M., Das, M. K. ve Ari, S., Automatic ECG arrhythmia classification using dual tree complex wavelet based features AEU - Int. J. Electron. Commun., 69, 4 (2015) 715–721.
195. Dilmac, S. ve Korurek, M., ECG heart beat classification method based on modified ABC algorithm Appl. Soft Comput. J., 36 (2015) 641–655.
196. Nazarahari, M., Ghorbanpour Namin, S., Davaie Markazi, A. H. ve Kabir Anaraki, A., A multi-wavelet optimization approach using similarity measures for electrocardiogram signal classification Biomed. Signal Process. Control, 20 (2015) 142–151.
197. Khalaf, A. F., Owis, M. I. ve Yassine, I. A., A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines Expert Syst. Appl., 42, 21 (2015) 8361–8368.
198. Escalona-Morán, M. A., Soriano, M. C., Fischer, I. ve Mirasso, C. R., Electrocardiogram classification using reservoir computing with logistic regression. IEEE J. Biomed. Heal. informatics, 19, 3 (2015) 892–8.
199. Oster, J., Behar, J., Sayadi, O., Nemati, S., Johnson, A. E. W. ve Clifford, G. D., Semisupervised ECG Ventricular Beat Classification With Novelty Detection Based on Switching Kalman Filters. IEEE Trans. Biomed. Eng., 62, 9 (2015) 2125–34.
200. Paul, B., Shanavaz, K. T. ve Mythili, P., A New Optimized Wavelet Transform for Heart Beat Classification J. Mech. Med. Biol., 15, 5 (2015) 1550081.

201. Alshraideh, H., Ootom, M., Al-Araida, A., Bawaneh, H. ve Bravo, J., A Web Based Cardiovascular Disease Detection System. J. Med. Syst., 39, 10 (2015) 122.
202. Das, M. K. ve Ari, S., Electrocardiogram Beat Classification Using S-Transform Based Feature Set J. Mech. Med. Biol., 14, 5 (2014) 1450066.
203. Sumathi, S., Beulah, H. L. ve Vanithamani, R., A wavelet transform based feature extraction and classification of cardiac disorder J. Med. Syst., 38, 9 (2014) 98.
204. Ceylan, R., Özbay, Y. ve Karlik, B., A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network Expert Syst. Appl., 36, 3 (2009) 6721–6726.
205. Melgani, F. ve Bazi, Y., Classification of electrocardiogram signals with support vector machines and particle swarm optimization. IEEE Trans. Inf. Technol. Biomed., 12, 5 (2008) 667–77.
206. Jiang, W. ve Kong, S. G., Block-based neural networks for personalized ECG signal classification. IEEE Trans. Neural Netw., 18, 6 (2007) 1750–61.
207. Güler, İ. ve Übeyli, E. D., ECG beat classifier designed by combined neural network model Pattern Recognit., 38, 2 (2005) 199–208.
208. Engin, M., ECG beat classification using neuro-fuzzy network Pattern Recognit. Lett., 25, 15 (2004) 1715–1722.
209. Osowski, S. ve Linh, T. H., ECG beat recognition using fuzzy hybrid neural network IEEE Trans. Biomed. Eng., 48, 11 (2001) 1265–71.
210. Kaya, Y. ve Pehlivan, H., Classification of Premature Ventricular Contraction in ECG Int. J. Adv. Comput. Sci. Appl., 6, 7 (2015) 34–40.
211. PhysioBank <http://www.physionet.org/physiobank/database/mitdb/> 1 Ocak 2016.
212. Goldberger, a. L., Amaral, L. a. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K. ve Stanley, H. E., PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals Circulation, 101, 23 (2000) e215–e220.
213. Nodes, T. ve Gallagher, N., Median filters: Some modifications and their properties IEEE Trans. Acoust., 30, 5 (1982) 739–746.
214. Kicmerova, D., *Methods for Detection and Classification in ECG Analysis*, Brno University of Technology, 2009.
215. Lee, G. K., Klarich, K. W., Grogan, M. ve Cha, Y.-M., Premature ventricular contraction-induced cardiomyopathy: a treatable condition. Circ. Arrhythm. Electrophysiol., 5, 1 (2012) 229–36.

216. Jie Zhou, Automatic detection of premature ventricular contraction using quantum neural networks, Third IEEE Symposium on Bioinformatics and Bioengineering, 2003. Proceedings., 2003, IEEE Comput. Soc, 169–173.
217. Kaya, Y., EKG'deki Normal ve Erken Karıncık Kasılması Vurularının Sınıflandırılması, Uluslararası 8. İstatistik Kongresi, Ekim, 2013, Antalya.



ÖZGEÇMİŞ

1979 yılında Adana'da doğan Yasin KAYA; ilköğretimini Kadirli Gafarlı Köyü İlkokulunda, ortaokulu Kadirli Merkez Ortaokulunda ve lise eğitimini Kadirli Endüstri Meslek Lisesi, Elektrik Bölümünde okul birincisi olarak tamamladı. 1995 yılında girdiği KTÜ Fen Edebiyat Fakültesi, İstatistik ve Bilgisayar Bilimleri Bölümü'nden 1999 yılında bölüm birinciliği derecesi ile mezun oldu. Aynı yıl KTÜ Enformatik Bölümünde Öğretim Görevlisi olarak çalışmaya başladı. 2000 yılında KTÜ Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı'nda yüksek lisans eğitimine başladı. İngilizce ve bilimsel hazırlık aşamalarını tamamladıktan sonra askerlik nedeni ile eğitimine ara verdi. 2003 yılında Hava Kuvvetleri Komutanlığında teğmen olarak vatani görevini tamamladı. 2006 yılında yüksek lisans eğitimini tamamlayarak, aynı yıl KTÜ Fen Bilimleri Enstitüsü, Elektronik Mühendisliği Anabilim dalında doktora eğitimine başladı. 2008 yılında aynı enstitüde bulunan Bilgisayar Mühendisliği Anabilim Dalı'na yatay geçiş yaptı. Çalışma hayatı boyunca Temel Bilişim Eğitimi konusunda dersler veren Yasin KAYA, bu konularda 4 adet kitapta ortak yazar olarak katkı sağlamış, halen KTÜ Uzaktan Eğitim Uygulama ve Araştırma Merkezinde Öğretim Görevlisi olarak çalışmakta ve evli olup bir çocuğu vardır. İyi derecede İngilizce bilmektedir.

Kaya, Y., Pehlivan, H. ve Tenekeci, M. E., Effective ECG beat classification using higher order statistic features and genetic feature selection, *Biomed. Res.*, 28, 17 (2017) 7594–7603.

Kaya, Y. ve Pehlivan, H., Classification of Premature Ventricular Contraction in ECG, *Int. J. Adv. Comput. Sci. Appl.*, 6, 7 (2015) 34–40.

Kaya, Y. ve Pehlivan, H., Feature Selection Using Genetic Algorithms for Premature Ventricular Contraction Classification, 9th International Conference on Electrical and Electronics Engineering, ELECO 2015, IEEE.

Kaya, Y. ve Pehlivan, H., Comparison of classification algorithms in classification of ECG beats by time series, 2015 23rd Signal Processing and Communications Applications Conference (SIU), 2015, IEEE, 407–410.