

**ÇOKLU SINIFLANDIRICI SİSTEMLERİ İLE KONJESTİF KALP YETMEZLİĞİ
TEŞHİSİ**

Ali NARİN

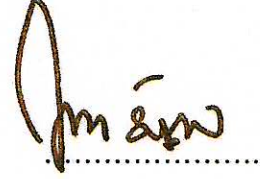
**Bülent Ecevit Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalında
Yüksek Lisans Tezi
Olarak Hazırlanmıştır**

**ZONGULDAK
Ağustos 2013**

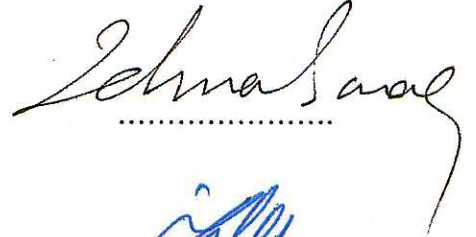
KABUL:

Ali NARİN tarafından hazırlanan "ÇOKLU SINIFLANDIRICI SİSTEMLERİ İLE KONJESTİF KALP YETMEZLİĞİ TEŞHİSİ" başlıklı bu çalışma jürimiz tarafından değerlendirilerek, Bülent Ecevit Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalında Yüksek Lisans Tezi olarak oybirliğiyle kabul edilmiştir.
28/08/2013

Başkan: Prof. Dr. Mahmut ÖZER (BEÜ)



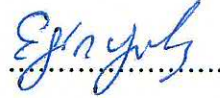
Üye : Doç. Dr. Zehra SARAÇ (BEÜ)



Üye : Yrd. Doç. Dr. Yalçın İŞLER (İKÇÜ)



Üye : Yrd. Doç. Dr. Ergin YILMAZ (BEÜ)



Üye : Yrd. Doç. Dr. İbrahim ALIŞKAN (BEÜ)



ONAY:


Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım .../.../2013



Prof. Dr. Şadi ŞEN

Fen Bilimleri Enstitüsü Müdürü V.

“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

Ali NARİN


ÖZET

Yüksek Lisans Tezi

ÇOKLU SINIFLANDIRICI SİSTEMLERİ İLE KONJESTİF KALP YETMEZLİĞİ HASTALARININ TEŞHİSİ

Ali NARİN

Bülent Ecevit Üniversitesi

Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Mahmut ÖZER

İkinci Danışman: Yrd. Doç. Dr. Yalçın İŞLER

Ağustos 2013, 81 sayfa

Kalbin temel görevleri arasında kanın vücuda pompalanması, metabolizma faaliyetleri sonucunda oluşan artık ürünlerin vücuttan uzaklaştırılması, vücut ısısının düzenlenmesi, asit-baz dengesinin korunması, hormon ve enzimlerin vücutun gerekli bölgelerine taşınması bulunmaktadır. Kalp Yetmezliği, kalbin kanla dolması veya kanı pompalaması yeteneklerinin azalması durumudur. Vücutun çeşitli yerlerinde sıvı birikmesi durumu bu rahatsızlığa sahip hastalarda yaygın olarak rastlandığı için, bu rahatsızlığa Konjestif Kalp Yetmezliği (KKY) ismi de verilmektedir. Kalp yetmezliğinin teşhisi uzman hekimler için kolay olmasına rağmen, hastalık belirtilerinin çoğu zaman diğer hastalıkların belirtileri ile karıştırılması nedeniyle özellikle pratisyen hekimler teşhiste zorlanmaktadırlar. Bu durumun ortadan kaldırılması için son zamanlarda veri madenciliği ve karar verme teknikleri oldukça gelişmiş ve bu doğrultuda yeni ileri teknikler sunulmuştur.

ÖZET (devam ediyor)

Bu çalışmada KKY hastalarının teşhisi için kalp hızı değişkenliği verileri üzerinden standart ölçümlerin yanında dalgacık dönüşümü ölçümlerinden ve doğrusal olmayan ölçümlerden oluşan özneliklerle çalışabilecek yeni bir teşhis sistemi amaçlanmıştır. Öncelikle farklı istatistiksel anlamlılık değerleri ve geriye doğru eleme yöntemi kullanılarak öznelik sayısını azaltma yoluna gidilmiştir. Azalan veri miktarı literatürde sık kullanılan 5 farklı sınıflandırıcı algoritması ile sınıflandırılmıştır. Bu sınıflandırıcılardan önce çalıştırılmak üzere, basitliği ve kolay uygulanabilir olması sebebiyle algılayıcılar eklenmiştir. Bu şekilde çok katlı sınıflandırıcılar tasarlanarak sistemin performans analizi yapılmıştır. Elde edilen sonuçlar doğrultusunda çok katlı sınıflandırma işlemi ile gerçekleşen bu yapının istatistiksel olarak öznelik sayısı düşürülmüş verileri kullanarak sınıflandırıcı algoritmalarında fark edilebilir bir iyileşme ve katlı sistemimizde de literatürdeki en yüksek sonuçlar elde edilmiştir.

Anahtar Kelimeler: Konjestif kalp yetmezliği, Kalp hızı değişkenliği, Öznelik seçimi, Örüntü tanıma

Bilim Kodu: 609.01.00

ABSTRACT

M. Sc. Thesis

DIAGNOSING THE PATIENTS WITH CONGESTIVE HEART FAILURE USING MULTI STAGE CLASSIFIER TECHNIQUES

Ali NARİN

**Bülent Ecevit University
Graduate School of Natural and Applied Sciences
Department of Electrical and Electronics Engineering**

Thesis Advisor: Prof. Mahmut ÖZER

Co-Advisor: Asst. Prof. Yalçın İŞLER

August 2013, 81 pages

The main purpose of the heart is to pump blood to the body, removal of waste products from the body as a result of metabolic activities, regulation of body temperature, protection of acid-base balance, transportation of the necessary parts of the body hormones and enzymes. Heart Failure is a decreased ability of the heart to either fill itself with blood or emptying it. The fluid accumulation in various parts of the body is common in the patients with heart failure, this disease is also named as Congestive Heart Failure (CHF). Although diagnosis of heart failure is easy for specialist physicians, some of physicians have difficulty in diagnosis. because symptoms of the disease are often mixed with other symptoms of the disease. Recently, to eliminate this situation, data mining and decision-making techniques are quite advanced and presented the new advanced techniques.

ABSTRACT (continued)

This study is focused on a new diagnostic system which using addition to the standard measurements is used wavelet transform measurement and nonlinear measurements on heart rate variability data for the diagnosis of CHF patients. In this system, the number of attributes is reduced by statistical method which different statistical significance values and backward elimination method. Decreasing the amount of data is classified with five different classification algorithms that commonly used in the literature. Perceptron is added due to its simplicity and easy-to-use features before five different classification algorithms. In this way multi stage classifier system is designed. Performance analysis of the system investigated. According to the results, the performance of the classifier algorithms is increased remarkably using features found statistical method which different statistical significance values and the highest results are obtained in our multi stage classifier system in the literature.

Key Words: Congestive heart failure, Heart rate variability, Feature selection, Pattern recognition

Science Code: 609.01.00

TEŐEKKÜR

Tez alıőmamın Danıőmanı olan Prof. Dr. Mahmut ÖZER'e alıőmamın her anında bana verdiği katkılardan ve desteklerden ötürü teőekkür ederim. Ayrıca alıőmalarımda bana yön gösteren, her konuda bana yardımcı olan tez alıőmamın İkinci Danıőmanı Yrd. Do. Dr. Yalın İŐLER'e verdiği katkı ve desteklerinden dolayı teőekkürü bir bor bilirim.

Tezin tüm aőamalarında maddi ve manevi desteklerinden dolayı oda arkadaşım Arő. Gör. Ferdi KARA'ya teőekkür ederim.

Ayrıca, bana olan desteklerinden dolayı manevi abim Okan TOĐAN'a da teőekkür ederim.

Son olarak, beni yetiőtiren, hayatım boyunca yanımda olan, sabırla eĐiten ve büyüten anneme ve babama teőekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL	ii
ÖZET	iii
ABSTRACT	v
TEŞEKKÜR	vii
İÇİNDEKİLER.....	ix
ÇİZELGELER DİZİNİ	xv
SİMGELER VE KISALTMALAR DİZİNİ.....	xvii
BÖLÜM 1 GİRİŞ	1
1.1 KONJESTİF KALP YETMEZLİĞİ	1
1.2 DOLAŞIM SİSTEMİ	2
1.3 KALP.....	3
1.3.1 Kalpteki Elektriksel İletim	4
1.3.2 Elektrokardiyografi	5
1.4 AMAÇ	7
BÖLÜM 2 KALP HIZI DEĞİŞKENLİĞİ ANALİZİ	9
2.1 GİRİŞ	9
2.2 FİZYOLOJİK ANLAMI	9
2.3 KLİNİK GEÇERLİLİĞİ.....	10
2.4 VERİ UZUNLUĞU SEÇİMİ	11
2.5 KALP HIZI VERİLERİNİN ELDE EDİLİŞİ.....	11
2.6 ÖNİŞLEM BASAMAKLARI	12
2.6.1 Hatalı verilerin yok edilmesi	13
2.6.2 Ara değerlendirme (interpolasyon).....	15

İÇİNDEKİLER (devam ediyor)

	<u>Sayfa</u>
2.6.3 Eğilim yok etme	18
2.7 KHD ANALİZİ ÖLÇÜMLERİ	19
2.7.1 Zaman alanı ölçümleri.....	19
2.7.2 Frekans alanı ölçümleri	20
2.7.2.1 Fourier Dönüşümü Yöntemi.....	21
2.7.2.2 Lomb-Scargle Periyodogram Yöntemi	22
2.7.2.3 Dalgacık Analizi Yöntemi.....	22
2.7.3 Doğrusal Olmayan Ölçümler	25
2.7.3.1 Poincare Çizimi	26
2.7.3.2 Yaklaşık Entropi	27
2.7.3.3 Eğilim Yok Edilmiş Dalgalanma Analizi.....	28
2.7.3.4 Örneklemeli Entropi	29
2.7.3.5 Sembolik Analiz.....	29
BÖLÜM 3 YÖNTEM	31
3.1 VERİLERİN ELDE EDİLMESİ.....	31
3.2 ÖN İŞLEMLER.....	32
3.3 ÖZİNİTELİKLERİN ÇIKARTILMASI.....	34
3.3.1 Özniteliklerin Ölçeklenmesi.....	37
3.4 ÖZİNİTELİK SEÇİMİ.....	38
3.4.1 Temel Bileşen Analizi	38
3.4.2 Geriye Doğru Eleme Yöntemi.....	39
3.4.3 Bağımsız Örneklem t-Testi	39
3.5 SINIFLANDIRICILAR.....	40
3.5.1 Algılayıcı Ağlar	40
3.5.2 Çok Katmanlı Algılayıcı Ağlar	41
3.5.3 Destek Vektör Makinaları	42
3.5.4 Radyal Tabanlı Fonksiyon Ağlar	43
3.5.5 K En Yakın Komşu	43

İÇİNDEKİLER (devam ediyor)

	<u>Sayfa</u>
3.5.6 Doğrusal Ayraç Analizi.....	44
3.5.7 K Ortalama Kümelemesi.....	45
3.6 GENETİK ALGORİTMA İLE SINIFLANDIRICI EĞİTİMİ.....	45
3.7 SINIFLANDIRICI BAŞARIMLARININ BELİRLENMESİ.....	46
3.7.1 Çapraz Geçerlilik.....	47
BÖLÜM 4 SONUÇLAR.....	49
4.1 TEMEL BİLEŞEN ANALİZİNİN TEŞHİSE ETKİSİ	49
4.2 K-ORTALAMA KÜMELEME ALGORİTMASININ TEŞHİS BAŞARIMI.....	51
4.3 FARKLI SINIFLANDIRICILARIN TEŞHİS BAŞARIMININ ANALİZİ	54
4.4 İSTATİSTİKSEL YÖNTEMLE ÖZNİTELİK SEÇİMİNİN TEŞHİSE ETKİSİ	55
4.5 ÇOK KATLI SINIFLANDIRICILARLA TEŞHİS BAŞARIMI.....	59
BÖLÜM 5 TARTIŞMA	65
KAYNAKLAR.....	67
ÖZGEÇMİŞ	81

ŞEKİLLER DİZİNİ

<u>No</u>	<u>Sayfa</u>
1.1 İnsanın dolaşım sistemi.....	3
1.2 Kalbin elektrofizyolojisi	5
1.3 Standart EKG ölçümü için elektrot bağlantısı	6
1.4 EKG göğüs derivasyonları bağlantısı	7
2.1 EKG'den KHD işaretinin elde edilişi	12
2.2 Vurular arasındaki sürelerin durumuna göre normal vuru, APC (kulakçık erken vurusu) ve VPC (karıncık erken vurusu)	14
2.3 RR verileri içinden karıncık kökenli ektopik vurunun tespiti örneği	14
2.4 Ara değerlendirme yöntemleri: solda doğrusal ve sağda kübik spline. İlk satırda verinin gerçek değerleri, ikinci satırda değerlere uydurulan fonksiyonları ve son satırda 10 Hz örnekleme hızında alınmış yeni veriyi göstermektedir	16
2.5 Kübik Spline ara değerlendirme yöntemi	17
2.6 Üç seviyeli dalgacık paket dönüşümü	23
2.7 Dalgacık paket dönüşümünün frekans alanı gösterimi	23
2.8 Daubechies 4 ana dalgacık fonksiyonu.....	24
2.9 SD1 ve SD2 ölçümleriyle birlikte örnek bir Poincare çizimi	26
3.1 KHD verileri	32
3.2 Yeniden örneklenmiş ve eğilimi yok edilmiş KHD verisi.....	33
3.3 Algılayıcı yapısı.....	40
3.4 Çok katmanlı algılayıcıların yapısı	41
3.5 SVM destek vektörleri ve aşırı düzlem.....	42
3.6 RBF yapay sinir ağının yapısı.....	43
3.7 KNN sınıflandırıcı örneği	44
3.8 İkiye ikilik karar matrisi	47
3.9 Geçerlilik hatası en düşük düzeyine ulaşıncaya kadar sınıflandırıcı eğitime devam edilir	48
4.1 K-Ortalama algoritması kullanarak farklı küme sayılarına göre elde edilen en yüksek başarımlar	50
4.2 Verinin yüklenmesi sonucu oluşan ara yüz görüntüsü	51
4.3 K-ortalama kümeleme algoritması ile kümeleme	52

ŞEKİLLER DİZİNİ (devam ediyor)

<u>No</u>		<u>Sayfa</u>
4.4	En yüksek başarımın elde edildiği 4 adet küme için görselleştirme ekranı.....	53
4.5	KKY minimum olduğu başlangıç değeri için diğer sınıflandırıcılardan geriye kalan kalan kişilerin birini hariç tut çapraz doğrulama yöntemiyle elde edilen sınıflandırıcı sistemi.....	60
4.6	KKY minimum olduğu başlangıç değeri ile tüm veri üzerinden diğer sınıflandırıcılardan geriye kalan kalan kişilerin test amaçlı kullanılması ile elde edilen sınıflandırıcı sistemi.....	60
4.7	Normal bireylerin minimum olduğu başlangıç değeri için diğer sınıflandırıcılardan geriye kalan kişilerin birini hariç tut çapraz doğrulama yöntemiyle elde edilen sınıflandırıcı sistemi.....	61
4.8	Normal bireylerin minimum olduğu başlangıç değeri için tüm veri üzerinden algılayıcı sınıflandırıcısından geriye kalan kalan kişilerin test amaçlı kullanılması ile elde edilen sınıflandırıcı sistemi	62

ÇİZELGELER DİZİNİ

<u>No</u>	<u>Sayfa</u>
3.1 Normal ve KKY hasta verilerine ait hasta bilgisi	34
3.2 Beş dakikalık veriler için elde edilen zaman alanı KHD analizi ölçümleri.....	35
3.3 Kısa süreli (5 dakikalık) veriler için elde edilen frekans alanı ölçümleri.....	35
3.4 Kısa süreli (5 dakikalık) veriler için doğrusal olmayan yöntemlerle elde edilen KHD ölçümleri	37
4.1 Temel bileşen analizi uygulama oranına göre ortalama sınıflandırıcı başarıları ve kullanılan öznelik sayıları	50
4.2 Farklı küme sayısı ve başlangıç değerleri için kümeleyici başarımları	53
4.3 5 farklı sınıflandırıcı ile sınıflandırma başarımları. En yüksek başarımlar (ACC), duyarlılık (SEN), özgünlük (SPE), pozitif seçicilik (POS) ve negatif seçicilik (NEG) .	55
4.4 Tüm öznelikler ve geriye doğru eleme yöntemi kullanılarak elde edilen öznelikleri kullanarak sınıflandırıcı performansının analizi, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznelik sayısını gösterir	56
4.5 %1, %2, %5, %10 ve %20 istatistiksel anlamlılık değerleri için öznelik seçimi ile sınıflandırıcı başarımları, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznelik sayısını gösterir	58
4.6 %1, %2, %5, %10 ve %20 istatistik anlamlılık değerleri için öznelik seçimi ile elde edilen öznelikler kullanılarak tüm veri üzerinden algılayıcı sınıflandırıcısından geriye kalan kalan kişilerin test amaçlı kullanılması ile elde edilen sınıflandırıcı başarımları, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznelik sayısını, KKY_min konjestif kalp yetmezliği hastalarının minimum hata ile sınıflandırılması, NORMAL_min normal bireylerin minimum hata ile sınıflandırılmış olduğu başlangıç değerlerini gösterir	63
4.7 %1, %2, %5, %10 ve %20 istatistik anlamlılık değerleri için öznelik seçimi ile elde edilen öznelikler kullanılarak algılayıcı sınıflandırıcısından geriye kalan kişilerden birini hariç tut çapraz doğrulama yöntemiyle elde edilen sınıflandırıcı başarımları, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznelik sayısını, KKY_min konjestif kalp yetmezliği hastalarının minimum hata ile sınıflandırılması, NORMAL_min normal bireylerin minimum hata ile sınıflandırılmış olduğu başlangıç değerlerini gösterir	64

SİMGELER VE KISALTMALAR DİZİNİ

- λ : Smoothness Priors yönteminde kullanılan düzenleme parametresi
RR_i : i-inci zamanda oluşan kalp vurusu ile bir önceki vuru arasında geçen süre
t_i : i-inci zaman değeri
T_i : i-inci zaman bir önceki zaman arasında geçen süre

KISALTMALAR

- ACC : Doğruluk
APC : Kulakçık erken vurusu
ApEn : Yaklaşık entropi (Approximate Entropy)
AR : Özyinelemeli (Auto Regressive)
AV : Atriyoventriküler düğüm
DFA : Eğilim yok edilmiş dalgalanma analizi
EKG : Elektrokardiyografi
FFT : Hızlı fourier dönüşümü
HF : Yüksek frekans
KHD : Kalp hızı değişkenliği
KKY : Konjestif kalp yetmezliği
KNN : k-En yakın komşu
LDA : Lineer diskriminant analizi
LF : Düşük frekans
LS : Lomb-Scargle
MLP : Çok katmanlı algılayıcı ağ
NEG : Negatif değerlilik
POS : Pozitif değerlilik
QRS : EKG içindeki karıncık kasılmasını gösteren kompleks işaret
RBF : Radyal tabanlı fonksiyon

SİMGELER VE KISALTMALAR DİZİNİ (devam ediyor)

- RR : Peş peşe gelen iki kalp vurusu arasında geçen süre
SA : Sinoatriyal düğüm
SampEn: Örnekleme entropisi
SEN : Hassaslık
SPE : Özgünlük
SVM : Destek vektör makinası
VLF : Çok düşük frekans
VPC : Karıncık erken vurusu
WE : Dalgacık enerjisi
WS : Dalgacık entropisi
WV : Dalgacık değışintisi (varyansı)

BÖLÜM 1

GİRİŞ

1.1 KONJESTİF KALP YETMEZLİĞİ

Kalp, vücudumuzun en önemli organlarından biridir. İhtiyaç duyulan besin maddelerini hücre ve dokulara taşıyan kanın pompalanması gibi çok önemli bir görevi vardır. Bu sebeple kalp vücudumuzda motor görevi yaparak kan dolaşımını yerine getirir (Berne and Levy 1997). Doku ve hücreler belli miktarda besin ve oksijene ihtiyaç duyarlar. Bu ihtiyaçlarının karşılanmaması yeterli miktarda kanın kalpten pompalanmamasına bağlıdır. Bu durumda kalp yetmezliği ortaya çıkar (Flavell and Stevenson 2001). Kalp yetmezliğinde kalp çalışmayı sürdürür fakat ihtiyaç duyulan miktar karşılanmadığından bu sorun daha başka rahatsızlıklara yol açabilir. Kan dolaşımındaki yetersizlikten dolayı kan dokularında ve damarlarda birikir. Vücudun çeşitli yerlerinde, örneğin kol, bacak ve akciğer gibi organlarda kan birikmesiyle konjestiflik meydana gelir. Kalp yetmezliği hastalarının çoğunda konjestiflik meydana geldiğinden bu rahatsızlık genellikle konjestif kalp yetmezliği (KKY) olarak adlandırılır (Wilbur and James 2005).

KKY son yıllarda adından sıkça söz ettirmektedir. Çünkü insanların ölüm sebeplerinden en önemlilerinden bir tanesidir (Albert 2000, Zambroski 2003). İstatistiklere göre ortalama yıllık 500,000 yeni KKY hastasının eklendiğini ve ortalama 250,000 KKY hastasının öldüğü uzmanlarca tahmin edilmektedir (Gura and Foreman 2004). KKY teşhisi konulan kişilerin yaşına bağlı olarak ölüm oranı değişmektedir. Genellikle teşhisten sonraki beş yıl çok önemlidir. Çünkü bu dönemde ölüm oranı hızı artmaktadır (Clark et al. 2003, Artinian et al. 2003). Amerikan kalp derneğinin bir araştırmasında beş yıl içerisinde KKY teşhisi konulan hastaların yaklaşık %50'si ölmektedir denilmiştir (American Heart Association 2006). Yukarıdaki istatistikler ABD'den elde edilmiş olmasına karşın aynı durumun ülkemiz içinde geçerli olduğu düşünülmektedir (İşler ve Kuntalp 2010). KKY rahatsızlığına erken teşhis konulabilirse hem ölüm oranı azaltılmış hem de kişilerin daha iyi yaşam sürmesi sağlanmış olur.

Kalp yetmezliđinin teŖhis edilmesi iin eŖitli yntemler vardır. İlk olarak kiŖinin klinik gemiŖi elde edilir (Shamsham and Mitchell 2000). Fakat tek baŖına bu yntem yeterli olmamaktadır. Bu yntem sadece yukarıda anlatılan belirtilerin teŖpit edilmesini sađlamaktadır (Wilbur and James 2005). Bu belirtiler de aslında tam olarak kalp yetmezliđine sahip kiŖileri teŖpit edememektedir. Bununla beraber kalp yetmezliđi teŖhisi iin izlenecek Framingham ve Boston gibi bazı teŖhis takip yntemleri mevcuttur. Fakat bu yntemlerin klinik olarak geerlilikleri hala sorgulanmaktadır (Di Bari et al. 2004).

Kalp yetmezliđinin teŖhis edilmesinde diđer yntemler elektrokardiyografi, ekokardiyografi ve gđs rntgeni baŖta olmak zere birok klinik inceleme testleri ve yntemleri vardır (Chiarugi et al. 2008). Bu yntemler sonucunda kalbin bir atım sonucunda ne kadar kanı vcuda pompaladıđı ve ne miktarda akciđerin oksijen sađladıđına bakılmaktadır. Bununla birlikte, hastaların fiziksel aktivite esnasında ne kadar oksijen sarf ettiđine bakılır. Buna efor testi denilir. Bu da kalp-akciđer egzersiz testleri ile mmkndr. zellikle efor testi ve ekokardiyografi yntemi kalp yetmezliđinin belirlenmesinde kullanılan nemli ltlerdir. Aslında bunların dıŖında kalp yetmezliđi rahatsızlıđına sahip kiŖilerin teŖpitinde belli bir yntem hala daha yoktur. zellikle bu rahatsızlıkla sıklıa karŖılaŖacak olan pratisyen hekimlerin rahatsızlıđı teŖpiti ancak rahatsızlıđın ilerlemiŖ aŖamalarında mmkn olabilmektedir. Bu durumda hastalık ilerlemiŖ ve hasta iin iinden ıkılmaz bir duruma yol aabilir.

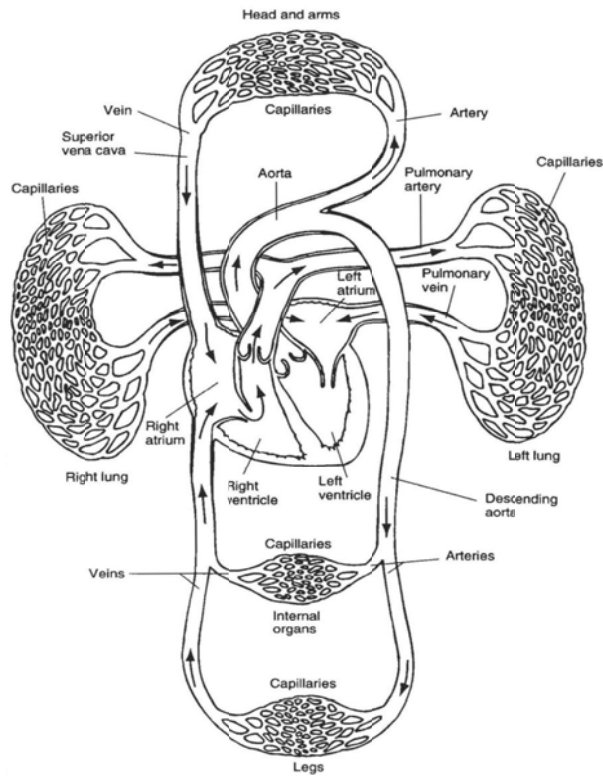
Avrupa'da gerekleŖtirilen bir araŖtırmaya gre, EKG kaydı iin baŖvuran adayların ortalama 48 saat sonrasına randevu verildiđi ve ekokardiyografi kaydı iin ise ortalama 3 ay sonraya randevu verildiđi teŖpit edilmiŖtir (Hobbs et al. 2005). Gnmzde bu srelerin azalmasına karŖın hala bazı blgelerde ve nfusun ok fazla olduđu yerlerde bu tr problemlerle karŖılaŖılmaktadır. Bu problemi ozmek ve kısa zaman ierisinde KKY rahatsızlıđını teŖhis etmek iin EKG tabanlı hızlı bir KKY teŖhis sistemine ihtiya duyulmaktadır.

1.2 DOLAŖIM SİSTEMİ

DolaŖım sisteminin merkezinde olan kalbin temel grevi vcutta kan akıŖını sađlamaktır. Bylece vcudun ihtiya duyuduđu besin ve oksijeni sađlamak, metabolizma faaliyetleri sonucu ortaya ıkan artıkları ve karbondioksiti vcuttan uzaklaŖtırmaktır (Narin vd. 2013a).

Bu şekilde olan dolaşım sistemi aslında kapalı döngü bir hidrolik sisteme benzemektedir (Webster 1998).

Genel olarak basit bir dolaşım sistemi Şekil 1.1’de gösterilmektedir. Sistemde görüldüğü gibi kalp pompa vazifesi görerek kanın vücuda dağılımını sağlar. Kanın kalpten vücuda dağılmasını atardamarlar sağlarken, kanın vücuttan toplanmasını ve tekrar kalbe geri gelmesini toplardamarlar sağlar. Atardamarlar ve toplardamarların oluşturduğu bu sistemde kan vücudun her yerine dağılır. Böylece ihtiyaç duyulan oksijen ve besin, doku ve hücrelere, atıkları ve karbondioksiti, akciğerlere ve böbreklere taşınmaktadır.



Şekil 1.1 İnsanın dolaşım sistemi (Webster 1998).

1.3 KALP

Kalp dışarıdan herhangi bir uyarıya ihtiyaç duymayan, kendi uyarısıyla kanı vücuda pompalayan özel çizgili kaslardan oluşmuş bir yapıdır. Kalp toplamda 2 bölüm ve 4 odacıktan oluşur. Üstteki iki odacığa kulakçık (atrium) ve alttaki iki odacığa ise karıncık (ventricle) adı verilir (Webster 1993). Kanın belli bir düzen içerisinde akışını sağlamak için iki tane kulakçık arasında ve iki tane de atardamar girişlerinde olmak üzere dört adet kapak

bulunmaktadır. Böylece kalbin kasılıp gevşemesiyle ve kapakların açılıp kapanması ile kan tek yönlü akışını tüm vücutta gerçekleştirir.

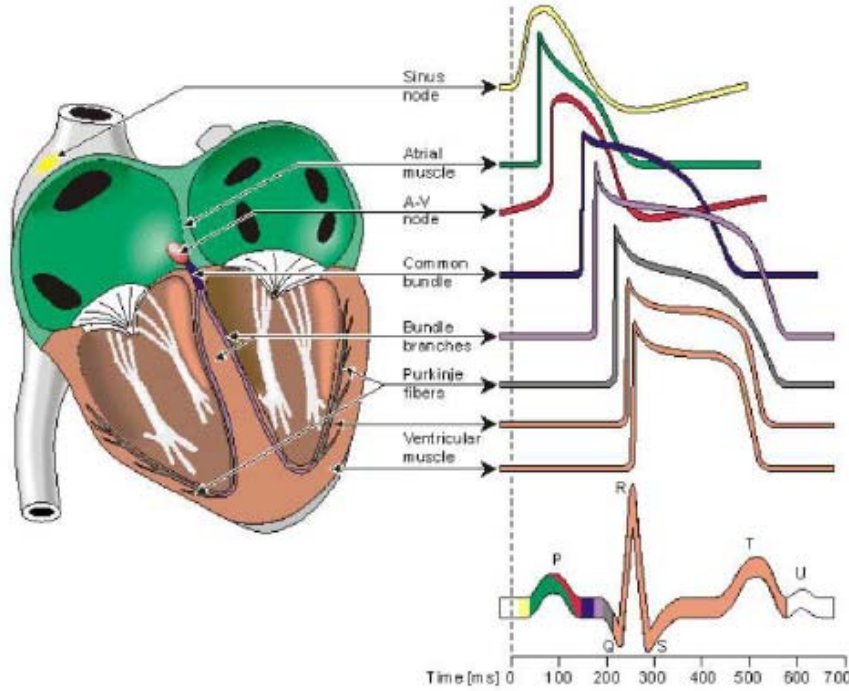
Kalp içindeki kanın kulakçık ve karıncıklardaki kasılma (sistol) ve gevşeme (diastol) elektriksel uyarımla gerçekleşmektedir. Öncelikle uyarımı alan kulakçıklar kasılır. Kulakçıkların kasılmasıyla orada bulunan kan kapakçıklar vasıtasıyla karıncıklara geçer. Karıncıklara geçen kan, belli bir miktar gecikme sonunda elektriksel uyarımı almasıyla karıncıklar kasılmaya başlar. Elektriksel uyarının bitmesiyle karıncıklar gevşer ve tekrardan kanla dolmaya başlar.

Kalpte gerçekleşen yukarıdaki olayda her iki kulakçık ve her iki karıncık birlikte kasılıp gevşer. Yani, sağ ve sol kulakçıklar aynı anda kasılmasıyla, diğer taraftan aynı anda sağ ve sol karıncıklarda gevşemektedir. Bu kasılıp gevşeme olayı birbirine zıt olarak gerçekleşmektedir. Sonuç olarak, kalbin kulakçıklarının gevşemesiyle akciğerden temiz kan sol kulakçığa ve vücuttaki kirli kan sağ kulakçığa gelir. Karıncıkların kasılmasıyla, kulakçıkların kasılması sonucu karıncığa gelen temiz kan sol karıncıktan vücuda pompalanırken kirli kan sağ karıncıktan akciğere temizlenmek için pompalanır. Böylece kan dolaşımı vücutta sağlanmış olur.

1.3.1 Kalpteki Elektriksel İletim

Kalbin düzenli ve kesintisiz bir şekilde kasılıp gevşemesi için kalp içinde özel bir iletim sistemi bulunmaktadır. Bu sistem, Şekil 1.2’de Sinoatrial (SA) Düğüm, Atrioventriküler (AV) Düğüm, His demeti ve Purkinje Lifleri olmak üzere dört bölümden meydana gelen kalbin elektriksel uyarı ve iletim sistemi görülmektedir. SA düğüm, kalbin kasılıp gevşemesini sağlayan ilk elektriksel uyarıyı üreten bir hücre topluluğudur. Bu sebeple uyarı odağında denilmektedir. SA düğümden çıkan elektriksel uyarı ilk olarak her iki kulakçığa yayılır, bunun sonucunda kulakçıklar hızlı bir şekilde kasılmaya başlar ve kulakçıklardaki kan karıncığa dolar. Elektriksel iletim buradan AV düğüme gelir. AV düğüm ise iletim görevini üstlenir. SA düğümün uyarı çıkartmaması durumlarında AV düğüm uyarı çıkartır. Böylece kesintisiz olarak kalbin çalışması sağlanmış olur. AV düğüme gelen uyarı kısa bir gecikme sonunda his demetine geçer buradan his demetinin sağ ve sol dallarında ikiye ayrılarak sağ ve sol karıncık kasındaki purkinje liflerine ulaşır. Purkinje liflerine gelen elektriksel uyarı karıncık kaslarına hızlı bir şekilde yayılarak karıncıkların kasılması sağlanır. Karıncıkların kasılmasıyla kan akciğer ve tüm vücuda pompalanmış olur.

Bu uyarı ve ileti sistemi bir bütün halinde çalışır. SA düğümünden çıkan uyarı kulakçıkların kasılmasını sağlarken, AV düğüm ve his demetlerin geçen ileti purkinje liflerine gelmesiyle de karıncıklar kasılır. Kulakçıklar kasılırken karıncıklar gevşer, karıncıklar kasılırken kulakçıklar gevşer. Bu şekilde uyarı ve iletim sistemik olarak gerçekleşmiş olur.



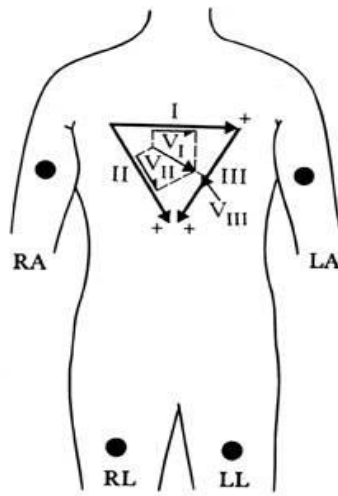
Şekil 1.2 Kalbin elektrofizyolojisi (Webster 1993).

1.3.2 Elektrokardiyografi

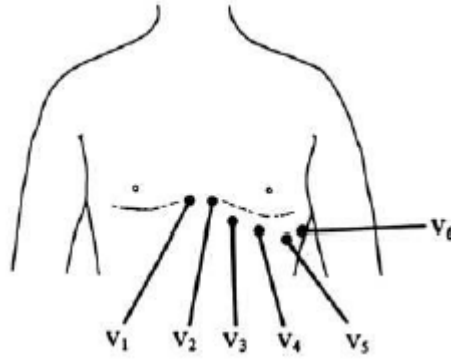
Kalp, SA düğüm ile elektriksel uyarıyı oluşturur ve kalbin her tarafına bu elektriksel uyarılar iletilmektedir. Bu sayede elektriksel uyarıyı alan kısımlar kasılma ve gevşeme işlemini gerçekleştirir. Kalpte oluşan ve iletilen elektriksel uyarılar aynı zamanda kalbi çevreleyen hücre ve dokulara da yayılır. Elektrokardiyografi ise birçok hücre ve dokunun elektriksel aktivitelerinin birleşiminden oluşmaktadır. Elektriksel uyarının azda olsa bir miktarı vücut yüzeyine ulaşmaktadır. Kalpteki bu elektriksel faaliyetler insan göğsünden veya kol ve bacaklarından ölçülebilecek düzeydedir. Kalpte oluşan elektriksel iletimi kaydetmek için kullanılan elektronik cihaza elektrokardiyograf, bu işlemin adına elektrokardiyografi ve kayıt sonunda elde edilen grafiğe ise elektrokardiyogram (EKG) denir. Şekil 1.2'de bu oluşan elektriksel dalga şekilleri ve bunların toplamı olan EKG işareti görülmektedir.

EKG işaretinin P, QRS ve T dalgalarından oluşan yapısı vardır (Şekil 1.2). P dalgası kulakçıkların kasılması esnasında oluşmaktadır. P-Q arasındaki zaman gecikmesi AV düğümdeki gecikmeden kaynaklanmaktadır. QRS dalga şekli ise karıncıkların kasılmasıyla oluşmaktadır. Ayrıca bu esnada kulakçıkların gevşemesi durumunda bulunmaktadır. Karıncıkları gevşemesi durumunda ise T dalga şekli oluşmaktadır. Sıkça olmasada bazı EKG dalga grafiklerinde küçük bir U dalgasına da rastlanmaktadır. Bu dalganın karıncıkların meme şeklinde olmasından kaynaklandığı düşünülmektedir (Conratha and Opthof 2005, Van Eck et al. 2005, Riera et al. 2008).

EKG işaretini elde ederken elektrot yüzeyleri arasındaki potansiyel farka bakılmaktadır. Elektrotların bağlanma şekilleride farklılık göstermektedir. Kol ve bacaklara veya göğse bağlanan elektrotlar yardımıyla potansiyel farklar ölçülür. Genellikle kullanılan bağlama şekli kol ve bacaklara bağlama yöntemidir (Şekil 1.3). Elektrotlar sağ kol (RA), sol kol (LA), sağ bacak (RL) ve sol bacağı (LL) olmak üzere 4 adet olacak şekilde bağlanır. Buradaki elektrotlar arasında ölçülen potansiyel farkları EKG deki lead adı verilen ölçümlere karşılık gelir. Örnek olarak Lead II için, negatif terminal elektrotu sağ kola ve pozitif terminal elektrotu sol bacağı takılır. Kalpteki QRS dalga şeklinin, yani karıncıkların kasılma anının en rahat bir şekilde Lead II içinde görülebildiği için, kardiovasküler sistemler içindeki rahatsızlıklar ve özellikle kalp hızı değişkenliği analizi ve ritim bozuklukları tespitinde daha çok Lead II kullanımı tercih edilir. Bununla beraber göğüs üzerine belli noktalara takılan elektrotlar sayesinde de EKG ölçümü yapılmaktadır (Şekil 1.4). Şekildeki V1, V2, V3, V4, V5 ve V6 ifadeleri göğüs derivasyonları olarak adlandırılır.



Şekil 1.3 Standard EKG ölçümü için elektrot bağlantısı (Webster 1998).



Şekil 1.4 EKG göğüs derivasyonları bağlantısı (Webster 1998).

EKG dalga şekli içerisinde QRS en yüksek genlikli bileşen olduğundan dolayı Kalp Hızı Değişkenliği (KHD) analizi QRS dalganın tespitine dayanmaktadır. QRS dalgasının tespiti için literatürde gerçekleştirilmiş birçok yöntem bulunmaktadır. Konuyla ilgili ayrıntılı bilgi için (Kohler et al. 2002) tarafından gerçekleştirilen çalışmalara bakınız.

1.4 AMAÇ

Son yıllarda kalbin peş peşe gelen vuru zamanları arasındaki değişimler üzerine yapılan çalışmalarda bir artış bulunmaktadır. Meydana gelen bu değişimler üzerinden yapılan çalışmalara genel olarak KHD adı verilmekte ve KKY hastalığının teşhisi üzerine gerçekleştirilen çalışmalar bulunmaktadır (Asyalı 2003, İşler ve Kuntalp 2007a, 2007b, Pecchia et al. 2011, Yu and Lee 2012a, 2012b, Jovic and Bogunovic 2011). Literatüre bakıldığında son zamanlarda, lineer ayraç analizi (LDA) ve Bayes sınıflandırıcıları kullanılarak bu konuda yapılmış Asyalı'nın çalışması bulunmaktadır. Asyalı yaptığı çalışmada 9 adet sık kullanılan uzun süreli (24 saat) zaman alanı ölçümleri ile FFT tabanlı frekans alanı ölçümlerinden çıkartılmış öznitelikler kullanılmıştır. Bu öznitelikler ile sınıflandırıcı başarımları olarak hassaslık ve özgünlük değerlerini sırasıyla 81,8% ve 98,1% olarak elde etmiştir (Asyalı 2003). İşler ve Kuntalp ise çalışmalarında standart zaman alanı ölçümlerini ve frekans alanı ölçümlerinin yanı sıra bunlara dalgacık entropi ölçümlerini de ekleyerek k-en yakın komşu sınıflandırıcısı ile 79.3% hassaslık ve 94.4% özgünlük değerlerini elde etmişler (İşler ve Kuntalp 2007a). Aynı yazarlar bir diğer çalışmalarında kalp hızı normalizasyon işlemi ekleyerek k- en yakın komşu sınıflandırıcısı ile 82.76% hassaslık ve 100% özgünlük değerlerini elde etmişler (İşler ve Kuntalp 2010). Pecchia ve arkadaşları, standard olmayan öznitelikler kullanarak Sınıflandırma ve Regresyon Ağacı sınıflandırıcısı ile 89.7% hassaslık ve 100% özgünlük değerlerini elde etmişler (Pecchia et al. 2011). Yu ve Lee

standard zaman ve bispektral analiz tabanlı frekans ölçümlerini kullanarak elde edilen öznitelikleri kullanarak RBF kernel fonksiyonlu destek vektör makinaları sınıflandırıcı algoritmasıyla ve genetik algoritmayla yapılan öznitelik seçimi sonucunda 95,55% hassaslık ve 100% özgünlük değerleri, doğrusal kernel fonksiyonlarını kullanarak ise 93,10% hassaslık ve 98,14% özgünlük değerlerini elde etmişler (Yu and Lee 2012a, 2012b). Jovic and Bogunovic, standard zaman alanı ve lineer olmayan ölçümler ile elde ettikleri öznitelikler ile SVM, MLP, C4.5 ve Bayes sınıflandırıcıları kullanarak sırasıyla 77.2%, 96.6%, 99.2%, ve 98.4% hassaslık değerleri ve sırasıyla 87.4%, 97.8%, 98.4% ve 99.2% özgünlük değerlerini elde etmişler (Jovic and Bogunovic 2011).

KKY teşhisi için kullanılabilecek birçok yöntem geliştirilmesine rağmen, bunların basit mikroişlemci tabanlı elektronik düzeneklerine aktarılmasında karşılaşılabilecek en büyük problem hafıza ve işlem hızı problemi olacaktır. Bu problemin aşılması için Öztürk tezinde öznitelik seçimi yapmadan örnek sayısını düşürerek k en yakın komşu sınıflandırıcısını başarımının artırılması çalışmasını yürütmüştür. Çalışmasında, Geriye doğru eleme, Eleme-Ekleme, Geriye doğru tekrarlı eleme ve Genetik algoritma yöntemleri kullanılarak elenecek örneklerin seçimini sağlamıştır. Biri-hariç (leave-one-out) yöntemi ile sınıflandırıcı başarımlarını tespit etmiştir. Sonuç olarak sınıflandırıcı başarımını k=1 değeri için %100 elde etmiştir (Öztürk 2012).

Bu çalışmada hafıza ve işlem hızı problemini çözmek için istatistiksel yöntemlerle ve geriye doğru eleme yöntemiyle öznitelik seçimi yapılmıştır. Elde edilen bu öznitelikler kullanılarak literatürde sık kullanılan sınıflandırma algoritmalarının performanslarının karşılaştırılması amaçlanmıştır. Üstelik çok katlı sınıflandırıcı sistemleri tasarlanarak basit algoritmalarla aşamalı olarak hasta ve normal bireylerden teşhisi kolay olanlar teşhis edilip, son aşamada daha güçlü olan algoritmalar kullanılarak teşhis başarımının artırılması hedeflenmiştir.

BÖLÜM 2

KALP HIZI DEĞİŞKENLİĞİ ANALİZİ

2.1 GİRİŞ

Kalp Hızı Değişkenliği (KHD) analizi son yıllarda üzerinde oldukça çok çalışılan bir konudur. KHD genel olarak kalbin her bir vurusu arasındaki süreyi ele almaktadır. Değişen bu süre kardiovasküler sistemdeki rahatsızlıkların ve kalp damar sağlığının önemli bir göstergesidir (Malik and Camm 1995). KHD'nin genel bir standardının olması ve klinik geçerliliğini ortaya koymak için ESC/NASPE Task Force (1996) adlı matematik, mühendislik, fizyoloji ve klinik tıp alanlarında çalışan toplam 13 kişiyle, KHD ile ilgili çalışmalara belli bir standart getirilmiştir. Bu bölümde genel olarak KHD hakkında bilgi, KHD analizi kullanılarak zaman ve frekans alanındaki özniteliklerin elde edilmesi ve doğrusal olmayan KHD ölçümleri hakkında bilgi verilecektir.

2.2 FİZYOLOJİK ANLAMI

KHD'nin oluşmasından sorumlu olan SA düğüm ile uyarımlar başlar. SA düğümdeki hücreler kendiliğinden aksiyon potansiyeli oluşturur. Bunun sonucunda kalbin her tarafına bu uyarımlar iletilir. Dinlenim durumunda olan kas hücreleri gelen uyarımlara göre kasılır. Kanın doldurma ve boşaltma işlemini gerçekleştirir. Kalp atım hızının, düzenli bir şekilde sağlanması merkezi sinir sisteminin sempatik ve parasempatik bölümleri tarafından kontrol edilir. KHD bu iki bölümün dolayısıyla otonom sinir sisteminin aslında bir göstergesidir. Parasempatik bölüm kalp atımını yavaşlatma yönünde etki yaparken, sempatik bölüm ise kalp atımını hızlandırma yönünde etki yapar (Chang et al. 2001). Bu şekilde, bu kontrol sistemi KHD'nin kaynağı ve belirleyicisi olarak kabul edilir (Cao 2004).

2.3 KLİNİK GEÇERLİLİĞİ

İnsan kalbi vuruları arasındaki deęişkenlięin klinik geęerlilięi ilk olarak 1965 yılında anne karnındaki bebeęin (fetusun) stres durumuna baęlı olarak kalp hızındaki deęişmelerin fark edilmesiyle ortaya çıkmıştır (Hon and Lee 1965). Daha sonra benzer şekilde (Sayers 1973) ve dięer arařtırmacılar (Hirsh and Bishop 1981, Luczak and Luring 1973, Penaz et al. 1968) kalp hızı iřareti içinde fizyolojik iřaretlerin varlıęını tespit etmişler. Ewing ve arkadaşları 1970'li yıllarda diyabet hastalarında otonomik nöropati tespiti için kısa süreli kalp hızı deęişimleri analizini kullanarak KHD'nin ne kadar etkili olduęunu ortaya koymuşlardır (Braune and Geisenorfer 1995, Ewing et al. 1985).

Bununla birlikte kalp krizi geęiren hastaların KHD deęeri düşük olmasından kaynaklanan ölüm oranının arttıęı ortaya çıkarılmıştır (Wolf et al. 1978). Bu ortaya konduktan sonra kalp krizi geęiren hastaların ölüm oranlarının tespitinde KHD'nin klinik olarak önemi artmıştır (Bigger et al. 1992, Huikuri et al. 2003, Kleiger et al. 1987, Lombardi et al. 2001, Malik et al. 1989, Malik et al. 1989). Kardiovasküler rahatsızlıklar ve buna baęlı dięer birçok rahatsızlıęın tespitiyle ilgili alıřmalar artmıştır.

Sonraki ařamalarda kalp hızı verileri ile güç spektral yoğunluęu analizlerinin nasıl yapılacaęı ortaya konmuştur (Akselrod et al. 1981). Bu alıřmalar ise KHD'nin otonom sinir sistemine olan iliřkisini frekans alanı boyutunda göstermesiyle oldukça etkili ve bu alanda sıkça kullanılmasında etkili olmuştur (Pagani 2000, Pagani et al. 1986, Pomeranz et al. 1985). 1996 yıllarında Task Force adlı bir grup bilim adamının alıřmalarıyla KHD ile ilgili klinik olarak belli standartlar belirlenmiştir (Task Force 1996).

Son yıllarda yukarıda anlatılanlar ışığı altında KHD ile ilgili çok miktarda yayın bulunmaktadır. Fakat KHD'nin klinik olarak geęerlilięini ortaya koyan iki alıřma vardır (Task Force 1996). Birincisi kalp krizi geęirmiş hastaların ölüm riskinin ve oranının ne olacaęının tespiti (Huikuri et al. 2003, Lombardi et al. 2001), ikincisi diyabetik nöropatinin tespit edilmesidir (Braune and Geisenorfer 1995, Pagani 2000). KHD ölçümlerinin genellikle cinsiyet, yař, fiziksel ve zihinsel durum, baskı ve dikkat yoğunluęu ile orantılı olduęu tespit edilmiştir. (Berntson et al. 1997). Bu kriterlerin de göz önüne alınması gerekmektedir.

2.4 VERİ UZUNLUĞU SEÇİMİ

Yapılan çalışmaların, literatürde yapılmış diğer çalışmalarla karşılaştırılması oldukça önemlidir. KHD ölçümlerin hesaplanması sırasında, veri uzunluğuna göre sonuçlar değişmektedir. Bu nedenle veri uzunluğunun belirlenmesi sonuçların güvenilirliği için çok önemlidir. Veri uzunluğu hastayı olumsuz duruma düşürmeyecek ve sıkıntı vermeyecek kadar kısa olmalıdır. Bununla birlikte kalpte gerçekleşen tüm fizyolojik ve patolojik değişimleri içerecek kadar uzun süreli olmalıdır. KHD yapı olarak durağan olmayan yapıdadır. Bu sebeple oldukça değişiklik göstermektedir. Onun için genellikle kısa süreli ölçümler tavsiye edilmektedir. Fakat kısa süreli ölçümlerde düşük frekanslı bileşenlerin olmaması bu ölçümlerin dezavantajıdır (Izard et al. 1991).

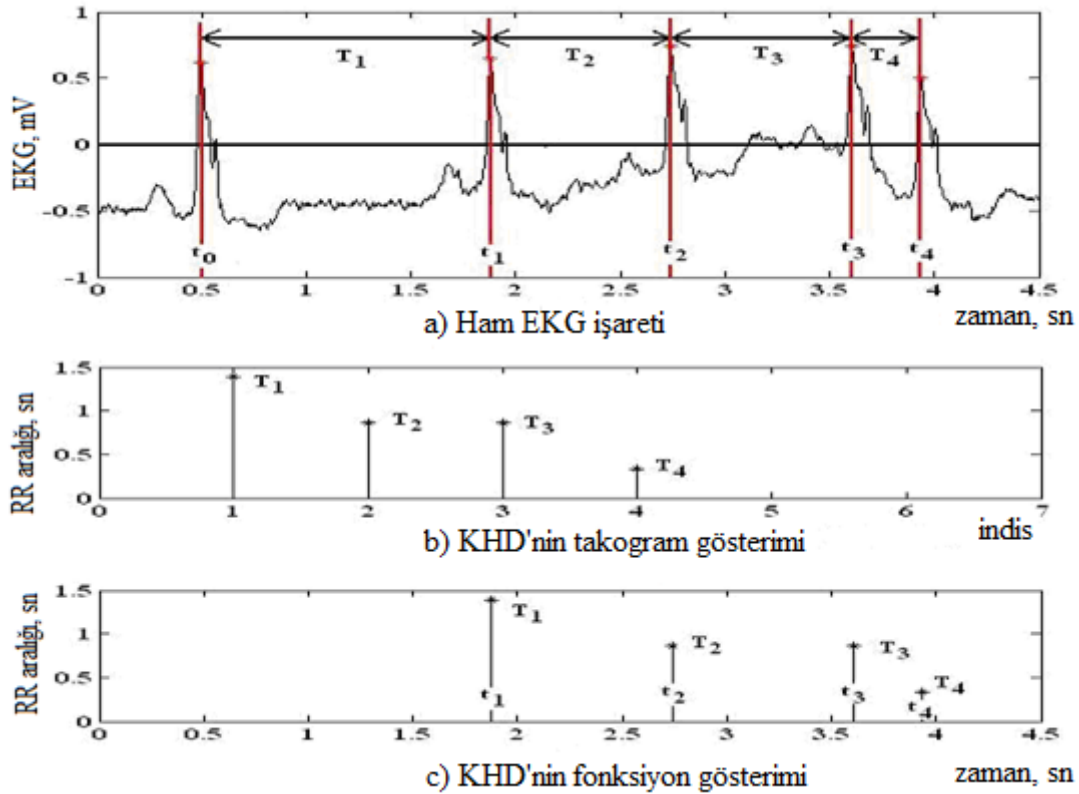
Literatürde, KHD'nin veri uzunluğuyla ilgili oldukça fazla çalışma vardır. Bunlar: 30 saniye (Izard et al. 1991), 1 dakika (Harper et al. 1976, Rosenstock et al. 1999, Schechtman et al. 1998), 2 dakika (Antila et al. 1990, Liao et al. 1996, Myers et al. 1986, Spicer and Lawrence 1987), 100 saniye (Stevens et al. 1988), 5 dakika (Berntson et al. 1997, Bigger et al. 1992, Edlinger et al. 1994, Kleiger et al. 1991, Rottman et al. 1990, Task Force 1996) ve 24 saatlik çalışmalardır (Leistner et al. 1980, Pikkujamsa et al. 1999). 1996 yılında Task Force adlı grub tarafından ortaya konan standartlara göre kısa süreli KHD ölçümleri için 5 dakikalık kayıt süresi ve uzun süreli KHD ölçümleri için ise 24 saatlik kayıt süresi tavsiye edilmektedir. Genel olarak, kısa süreli kayıtlar sempatik ve parasempatik bölgelerin değişimlerinin etkilerinde ve uzun süreli kayıtlar ise kişinin günlük fiillerinin sinir sistemindeki değişimlerinin incelenmesinde kullanılmaktadır.

Güvenilirlik ve tekrarlanabilirlik noktasından kısa süreli KHD ölçümleri literatürde sıkça kullanılmıştır. Kullanılabilir olduğu da bazı çalışmalarda ifade edilmiştir (Dionne et al. 2002, Hamilton et al. 2004, McNames and Aboy 2006). Yapılan bu çalışma ise hem elde edilmesi daha kolay hemde güvenilir sonuçlar vermesi sebebiyle EKG'den kısa süreli (5 dakikalık) KHD ölçümleri alınmıştır.

2.5 KALP HIZI VERİLERİNİN ELDE EDİLİŞİ

KHD verileri genel olarak daha öncede sözü edilen EKG işaretindeki QRS dalga yapısının belirlenmesi sonucunda elde edilir (İşler 2009). Çünkü QRS dalga yapısı EKG işareti

içerisinde en belirgin genlik değerine sahip bileşendir. QRS tespit edildikten sonra peş peşe gelen RR zaman aralıkları arasındaki zaman farkı (yani, $T_n = t_n - t_{n-1}$) olarak tanımlanır (Şekil 2.1(a)). Tüm veriler üzerinde bu işlem yapıldıktan sonra, (n, T_n) veri çifti olarak takogram şeklinde gösterilebilirken (Şekil 2.1(b)) ya da (t_n, T_n) veri çifti olarak fonksiyon şeklinde (Şekil 2.1(c)) gösterilebilir. Fakat literatürde genellikle fizyolojik açıdan daha anlamlı olduğu için fonksiyon şeklinde gösterimi tercih edilmektedir (Task Force 1996).



Şekil 2.1 EKG'den KHD işaretinin elde edilişi (İşler 2009).

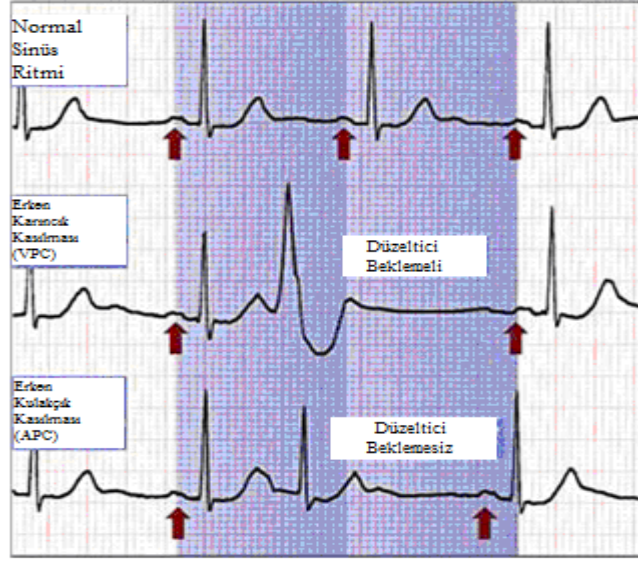
2.6 ÖNİŞLEM BASAMAKLARI

Önişlem basamaklarında, üzerinde çalışılan EKG işaretinden KHD ölçümleri yapabilmek için EKG verisinin hatalı vuru içermemesi, durağan olması ve zamanda eşit aralıklarla örneklenmiş olması gerekmektedir. (Lynn and Chiang 2001). EKG verisinin yukarıdaki özellikleri sağlaması için takip eden konu başlıklarında yapılacak işlemlerin nasıl ve hangi yöntemlerle yapıldığı anlatılacaktır.

2.6.1 Hatalı verilerin yok edilmesi

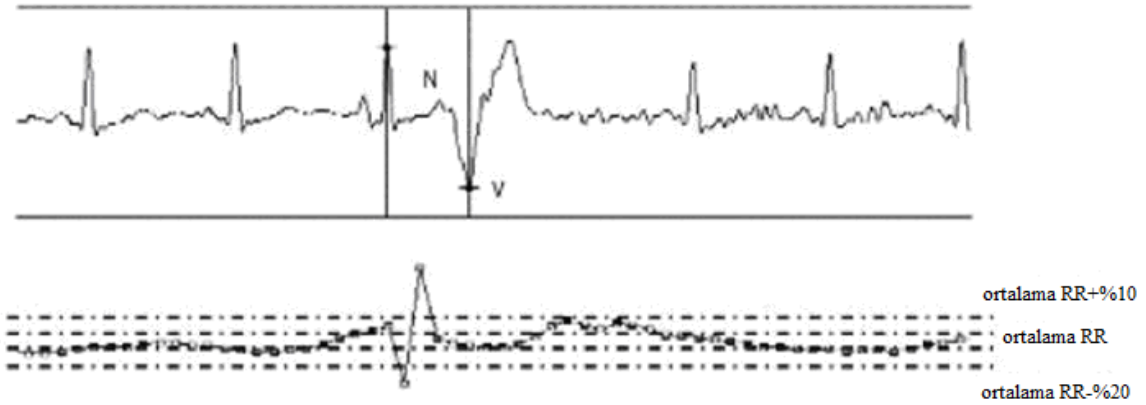
KHD ölçümünün düzgün bir şekilde elde edilmesi için RR zaman aralıkları içindeki herhangi bir hatalı verinin olmaması gerekmektedir (Clifford et al. 2006). RR zaman aralıkları arasındaki hatalı verilerin kaynağı iki türdür. Bunlar, uygulanan teknik ve kişinin fizyolojik durumundan kaynaklıdır. QRS yapısını tespiti sırasında oluşan hatalı veriler teknik kaynaklıdır. Bu hatalı verilerin kaynağı kullanılan tespit algoritmasının ölçüm hatalarından veya hesaplamada yaptığı yanlışlıktan kaynaklanır. Diğer hatalı veri kaynağı olan fizyolojik hatalı veriler ise ektopik (olması gerekenden farklı kaynaktan oluşan) vurular ve aritmik (ritim bozukluğu olan) olayları içerir. Hatalı olan bu verilerden kurtulmak için verilerin EKG uzmanına gösterilmesi ve tespit edilmesi gerekmektedir. Böylece en az hatalı veri içeren EKG verisi kullanılması önerilmektedir (Task Force 1996). Bu işlem yapıldığında, elde çok fazla veri kalmamışsa o zaman bu hatalı verilerin yerine uygun bir ara değerlendirme yöntemiyle (interpolasyon) değer üretilir ve hatalı verilerin sonucunda çıkacak olan olumsuzluklar ortadan kaldırılmış olur (Clifford and Tarassenko 2005, Lippman et al. 1993, 1994, Mateo and Laguna 2003).

Fizyolojik hatalı veriler için ilk olarak SA düğümünden kaynaklanmayan QRS dalga yapısının tespit edilmesi çok önemlidir (Clifford et al. 2006). Daha sonra kulakçık ve karıncıkların erken kasılması (APC ve VPC) RR takogram gösteriminde kolayca tespit edilebilir. Yani, erken vuru oluşumunda takogram değerinde bir düşüş meydana gelir. Vurunun kulakçık erken vurusu mu (APC) yoksa karıncık erken vurusu mu (VPC) olduğunu anlamak için takip eden bir sonraki vuruya bakılır. Şekil 2.2’de görüldüğü gibi eğer vuru zamanı ortalama vuru süresi kadarsa APC, daha uzunsa VPC olarak adlandırılır



Şekil 2.2 Vurular arasındaki sürelerin durumuna göre normal vuru, APC (kulakçık erken vurusu) ve VPC (karıncık erken vurusu) (Clifford et al. 2006).

Literatürdeki bir algoritmaya göre, ektopik vurular ortalama RR değerinin %20'den fazla aşağı düşen vuruları için tanımlanmaktadır (Langley et al. 2001). Bu çalışmaya göre ektopik vuruşlardan sonra gelen vuru ortalama RR süresinden $\pm 10\%$ içinde oluşursa kulakçık kökenli APC vurusu ve $\pm 30\%$ 'dan daha uzun sürede oluşursa karıncık kökenli VPC vurusu olduğu sonucuna ulaşılır (Şekil 2.3).



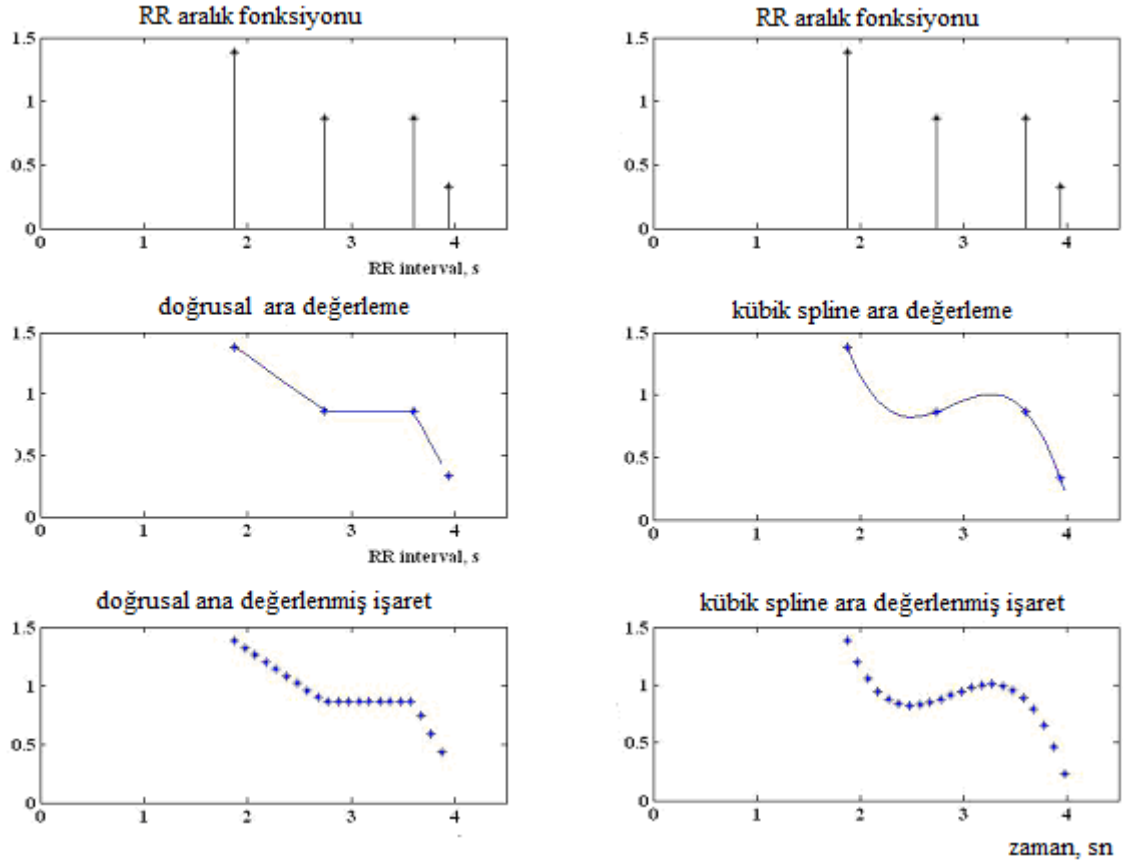
Şekil 2.3 RR verileri içinden karıncık kökenli ektopik vurunun tespiti örneği (Langley et al. 2001).

KHD ölçümlerinde kulakçık tabanlı ve karıncık tabanlı erken vuruların ölçüm sonuçlarını etkilediği anlatılmıştır (Clifford and Tarassenko 2005). Bu nedenle, KHD ölçümlerinde ektopik vuruşlardan arındırılması gerekmektedir. Daha sonra ektopik verilerden arındırılmış verilerin yerine ara değerlendirme yöntemleri kullanılarak veri kaybı önlenmiş ve daha sağlıklı KHD ölçümleri elde edilebilmektedir.

2.6.2 Ara deęerleme (interpolasyon)

Fourier donüşümü gibi bazı frekans alanı ölçümleri zamanda eşit aralıklarla örneklelenmiş verilerde gerçekleştirilir. Oysaki KHD verileri zamanda eşit aralıklarla örneklelenmiş veriler değildir. Bu nedenle bu verilerin ara deęerleme (interpolasyon) yöntemleri ile yeniden örneklelenmesi gerekmektedir. Yeniden örnekleme yönteminde en önemli ölçüt örnekleme frekansının doğru seçilmesidir. Bu yüzden örnekleme frekansının girişime sebep olmaması için işaretin frekansından en az iki kat daha büyük seçilmelidir. Üzerinde çalışılan KHD ölçümleri için en yüksek frekans bileşeni 0.5 Hz olarak belirtilmiştir. Fakat bazı fiziksel durumlardan sonra bu deęer daha yüksek çıkabiliyor olmasından dolayı örnekleme frekansını 4 Hz olarak seçilmesi uygun görülmüştür (Berntson et al. 1997). Fakat dięer bazı çalışmalarda yüksek frekanslı bileşenlerin olumsuz olarak etkilenmemesi için örnekleme frekansının 7 Hz ve üzeri olarak tercih edilmektedir (Clifford and Tarassenko 2005). Yaptığımız bu çalışmada yeterli olmasından dolayı 4 Hz örnekleme frekansı tercih edilmiştir.

Ara deęerleme yöntemleri arasında parçalar arası sürekliliğin ve doğrusal ara deęerlendirmeye göre geçişlerinin keskin olmaması nedeniyle kübik spline yöntemi oldukça kullanışlıdır (Kreyzig 1993). Doğrusal ve kübik spline ara deęerleme yöntemleri Şekil 2.4'te gösterilmiştir (İşler 2009).



Şekil 2.4 Ara değerlendirme yöntemleri: solda doğrusal ve sağda kübik spline. İlk satırda verinin gerçek değerleri, ikinci satırda değerlere uydurulan fonksiyonları ve son satırda 10 Hz örnekleme hızında alınmış yeni veriyi göstermektedir (İşler 2009).

Matematiksel olarak kübik spline ara değerlendirme yönteminde $N + 1$ ayrı verilen $a = x_0 < x_1 < \dots < x_{N-1} < x_N = b$ şartını sağlayacak şekilde x_i değerine karşılık gelen $y_i = f(x_i)$ için N tane üçüncü dereceden denklem yazılabilir:

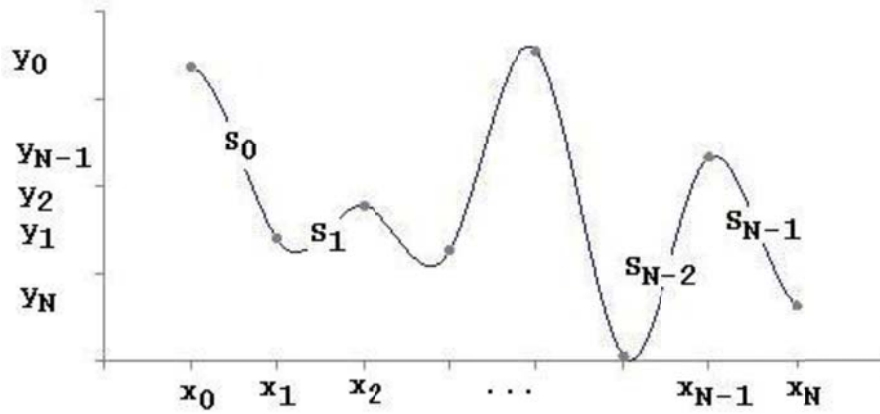
$$S(x) = \left\{ \begin{array}{ll} S_0(x) & x \in [x_0 \quad x_1] \\ S_1(x) & x \in [x_1 \quad x_2] \\ \vdots & \vdots \\ S_{N-1}(x) & x \in [x_{N-1} \quad x_N] \end{array} \right\} \quad (2.1)$$

burada

$$S_i(x) = \sum_{k=0}^3 (\beta_{i,k} (x - x_i)^k) \quad (2.2)$$

x ve y değerlerinden elde edilen fonksiyon Şekil 2.5'te gösterilmiştir. Aşağıdaki koşulların sağlanması halinde $S(x)$ spline fonksiyonu yazılabilir.

- Ara değerlendirme özelliği: $S_i(x_i) = f(x_i)$, $i = 0, 1, \dots, N$,
- Fonksiyonların kesiştiği x değerlerinde aldığı değerlerin eşit olması özelliği: $S_{i-1}(x_i) = S_i(x_i)$, $i = 1, 2, \dots, N - 1$,
- Bu noktalarda iki kez türevlenebilir olması özelliği: $S'_{i-1}(x_i) = S'_i(x_i)$ ve $S''_{i-1}(x_i) = S''_i(x_i)$, $i = 1, 2, \dots, N - 1$.



Şekil 2.5 Kübik Spline ara değerlendirme yöntemi (İşler 2009).

N tane kübik polinomun belirlenebilmesi için $4N$ tane denkleme gerek vardır. Aslında $4N-2$ adet denkleme rahat bir şekilde ulaşılmaktadır. İlk koşuldan $N + 1$ denklem, ikinci koşuldan $N-1$ denklem ve üçüncü koşuldan $2N-2$ denklem toplamda $4N-2$ denklem oluşmuş olmaktadır. Bunların dışında iki tane denkleme ihtiyaç vardır. Bunun için literatürde tavsiye edilen iki metot vardır: ilk olarak kenetlemeli kübik spline ($S'(x_0) = u$ ve $S'(x_n) = v$) ve ikinci olarak doğal kübik spline ($S''(x_0) = 0$ ve $S''(x_n) = 0$) metotlarıdır. Aşağıda doğal kübik spline denklemleri tanımlanmaktadır (Schilling and Sandra 1999):

$$S_i(x) = \frac{z_{i+1}(x - x_i)^3 + z_i(x_{i+1} - x)^3}{6h_i} + \left(\frac{y_{i+1}}{h_i} - \frac{h_i}{6} z_{i+1} \right) (x - x_i) + \left(\frac{y_i}{h_i} - \frac{h_i}{6} z_i \right) (x_i - x) \quad (2.3)$$

burada

$$h_i = x_{i+1} - x_i \quad (2.4)$$

ve

$$\begin{aligned} z_0 &= 0 \\ h_{i-1}z_{i-1} + 2(h_{i-1} + h_i)z_i + h_i z_{i+1} &= 6 \left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \right) \\ z_n &= 0 \end{aligned} \quad (2.5)$$

olmaktadır.

2.6.3 Eğilim yok etme

KHD ölçümleri için eğilim yok etme çok önemlidir. Çünkü KHD verilerinin yapısı gereği durağan değildir (Berntson et al. 1997). Bu durum KHD analizini çok etkilemektedir. Bu etkilenmenin en aza indirgenmesi oldukça önemlidir. Bu sebeple analizden önce eğilimler yok edilmelidir. Bu tür durağan olmayan ve analizi kötü etkileyen veri parçalarının analize dâhil edilmemesi (Weber et al. 1992, Grossman 1992) veya eğilim yok etme yöntemlerinin kullanılması tavsiye edilmektedir. Eğilimlerin yok edilmesiyle ilgili birçok yöntem vardır. Bunlardan bazıları, süzgeç kullanılması (Berntson et al. 1997), birinci dereceden model (Litvack et al. 1995, Mitov 1998) yüksek dereceden polinomsal model uydurulması (Mitov 1998) ve Smoothness Priors yöntemi (Tarvainen et al. 2002) olarak ifade edilebilir. Bunlara ek olarak durağan olmayan işaretlerde de ölçüm yapmayı sağlayan Dalgacık Analizi gibi yöntemler bulunmaktadır.

Smoothness Priors yöntemi en çok kullanılan eğilim yok etme yöntemlerinden biridir. Çünkü uygulaması kolaydır (Litvack et al. 1995). Bununla beraber matlab kodlarında mevcuttur. Smoothness Priors yönteminde örnek olarak KHD verisinin kullanılması (Tarvainen et al. 2002) KHD alanında çalışanları bu yönetime yöneltmiştir. Yöntemin diğer kolaylıkları arasında frekans cevabı tek bir parametre (λ) ile değiştirilebilmektedir. Tüm bu sebeplerden dolayı çalışmada bu yöntem tercih edilmiştir. Düzenleme parametresi için 1000 değeri kullanılmış ve ikinci dereceden türev ifadeleri kullanılarak durağan veriler aşağıdaki formülle ifade edilmiştir:

$$x_{durağan} = x - H\hat{\theta}_\lambda = \left(I - (I + \lambda D_2^T D_2)^{-1} \right) x \quad (2.6)$$

Burada x kalp hızı değişkenliği verisini, D_2 en iyi kalp hızı değişkenliği verisi eğimini, λ düzenleme parametresini, $x_{durağan}$ dalgalanması yok edilmiş işareti ifade etmektedir. Detaylı bilgi için ilgili atıflara bakılabilir (Litvack et al. 1995, Tarvainen et al. 2002).

2.7 KHD ANALİZİ ÖLÇÜMLERİ

Bu bölümde Task Force (1996) tarafından standart hale getirilmiş kısa süreli KHD ölçümleri ve literatür de sıkça kullanılan kısa süreli KHD ölçümleri ele alınacaktır. KHD ölçümleri zaman alanı, frekans alanı ve doğrusal olmayan ölçümler olmak üzere genel olarak 3 kısma ayrılmaktadır. Sonraki bölümlerde bu 3 ölçümle ilgili yöntemler açıklanacaktır.

2.7.1 Zaman alanı ölçümleri

Zaman alanı ölçümleri, standart KHD ölçümleri içerisinde yapılması en kolay ve pratik ölçümlerdir. Bu ölçümlerde kalbin her bir vurusu arasındaki yani kalp hızı değişkenliği verilerinin istatistiksel olarak ele alınmasıyla ölçümler yapılmaktadır

Zaman alanı ölçümleri içerisinde, KHD verilerinden elde edilen ortalama, standart sapma, peş peşe gelen kalp hızı verileri arasındaki farkların standart sapması ve RMS değeri, peş peşe gelen iki kalp hızı arasındaki farkı 20 ms ve 50 ms'den fazla olanların sayıları ve oranları gibi ölçümler vardır (Bigger et al. 1989, Antila et al. 1990, Izard et al. 1991, Malik and Camm 1995, Berntson et al. 1997, Mietus et al. 2002).

Bu çalışmada ortalama, R-R aralıklarının standart sapması, peş peşe gelen kalp hızı verileri arasındaki farkların standart sapması ve RMS değeri olmak üzere 4 ölçüm yapılmıştır. Sırasıyla formülleri:

$$ORTALAMA = \frac{1}{N} \sum_{i=1}^N RR_i \quad (2.7)$$

$$SSRR = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (RR_i - ORTALAMA)^2} \quad (2.8)$$

$$SSPPRR = \sqrt{\frac{1}{N-1} \sum_{i=2}^N \left(RR_i - RR_{i-1} - \frac{1}{N-1} \sum_{k=2}^N (RR_k - RR_{k-1}) \right)^2} \quad (2.9)$$

$$RMS = \sqrt{\frac{1}{N-1} \sum_{i=2}^N (RR_i - RR_{i-1})^2} \quad (2.10)$$

Burada RR_i i. peş peşe gelen R-R aralığı süresini ve N değeri R-R aralığı sayısını ifade etmektedir.

2.7.2 Frekans alanı ölçümleri

Vücuttaki sempatik ve parasempatik değişimlerin etkilerini anlamaya yarayan frekans alanı ölçümleri KHD ölçümlerinde önemli bir yer tutar. Literatürde, frekans alanı ölçümleri kullanılan 3 frekans bandı vardır (Bigger et al. 1992, Costa et al. 1994, Kleiger et al. 1991, Mazursky et al. 1998, Mrowka et al. 1996, Myers et al. 1986, Ori et al. 1992, Task Force 1996). Bunlar, Task Force adlı bir grubun 1996 yılında yayınladıkları çalışmada çok düşük frekans (VLF) (0-0.04 Hz), düşük frekans (LF) (0.04-0.15 Hz) ve yüksek frekans (HF) (0.15-0.40 Hz) olmak üzere üçe ayrılmıştır (Task Force 1996).

Bu frekans bantlarının fizyolojik olarak anlamlı olduklarını söyleyen literatürde çalışmalar bulunmaktadır. Çok düşük frekans bileşenleri (VLF) gündelik yaşam içinde vücut için gerekli faaliyet düzenini (Barrett et al. 2001, Braga et al. 2002) ve ısı ve hormonal sistemlerin (Braga et al. 2002, Porter and Rivkees 2004, Vornanen et al. 2002, Williams et al. 2002) göstergesinde kullanılmaktadır. Düşük frekans bileşeni (LF), sempatik ve parasempatik aktivitenin göstergesinde (Goldstein et al. 1998, Lanfranchi and Somers 2002, Malpas 2002) ve yüksek frekans (HF) bileşenide parasempatik aktivite ve solunumla ilgili oluşan sinüs aritmisi rahatsızlıklarının göstergesinde (Barbieri et al. 2002, Rentero et al. 2002) kullanılmaktadır.

KHD ölçümlerinde frekans bileşenlerini hesaplamak için FFT, AR model, Lomb ve Dalgacık analizi tabanlı analiz yöntemleri kullanılmaktadır. Frekans bileşenlerine karşılık gelen frekans alanı ölçümleri: Toplam güç, VLF frekans bandı gücü, LF frekans bandı gücü, HF frekans bandı gücü, Normalize edilmiş LF bandı gücü (LF bandı gücünün LF ve HF bantlarının toplamı içindeki oranı), Normalize edilmiş HF bandı gücü (HF bandı gücünün LF ve HF bantlarının toplamı içindeki oranı) ve LF/HF oranı kullanılmaktadır (Bigger et al. 1989, 1992, Rottman et al. 1990, Kleiger et al. 1991, Ori et al. 1992, Task Force 1996, Yeragani et al. 1998).

2.7.2.1 Fourier Dönüşümü Yöntemi

N adet eşit zaman aralıklarıyla örneklenmiş ve zaman alanında ortalama ve varyans değeri değişmeyen x işaretinin Ayrık Fourier Dönüşümü (AFD) şu şekilde bulunabilir:

$$X(k) = \sum_{n=0}^{N-1} \left(x(n) e^{-j2\pi kn/N} \right) \quad (2.11)$$

burada $k = 0, \bar{1}, \bar{2}, \dots, \bar{(N-1)}$.

(2.11) denklemindeki N değerinin yüksek seçilmesi AFD hesaplamasını yavaşlatmakta ve verimsiz hale getirmektedir. Bu soruna çözüm olarak, hesaplamayı daha hızlı bir şekilde yapan ve aynı sonucu veren Hızlı Fourier Dönüşümü (FFT) yöntemi kullanılmaktadır (Cooley and Tukey 1965). FFT yönteminin kullanımı ve hesapsal olarak nasıl yapıldığı konuyla ilgili birçok kitapta vardır. O kitaplara bakılabilir (Ingle and Proakis 2000).

Parseval ilişkisine göre frekans alanı enerji değerleri hesaplanabilir:

$$E_x = \sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2 \quad (2.12)$$

Buradaki $|X(k)|^2/N$ ifadesi periyodik veri seti için güç spektral yoğunluğunu ifade etmektedir.

2.7.2.2 Lomb-Scargle Periyodogram Yöntemi

Lomb-Scargle Periyodogram Yöntemi (LS), yeniden örnekleme ve eğilim yok etmeye gerek kalmadan güç spektral yoğunluğu hesabının yapıldığı yöntemdir (Clifford and Tarassenko 2005, Laguna et al. 1998, Lomb 1976, Scargle 1982). Bu yöntemde kullanılan matematiksel formül:

$$P(w_n) = \frac{1}{2\sigma^2} \left\{ \frac{\left[\sum_{j=0}^{N-1} (x(t_j) - \bar{x}) \cos(w_n(t_j - \tau)) \right]^2}{\sum_{j=0}^{N-1} \cos^2(w_n(t_j - \tau))} + \frac{\left[\sum_{j=0}^{N-1} (x(t_j) - \bar{x}) \sin(w_n(t_j - \tau)) \right]^2}{\sum_{j=0}^{N-1} \sin^2(w_n(t_j - \tau))} \right\} \quad (2.13)$$

burada

$$\tau \equiv \frac{1}{2w} \tan^{-1} \left(\frac{\sum_{j=1}^N \sin(wt_j)}{\sum_{j=1}^N \cos(wt_j)} \right) \quad (2.14)$$

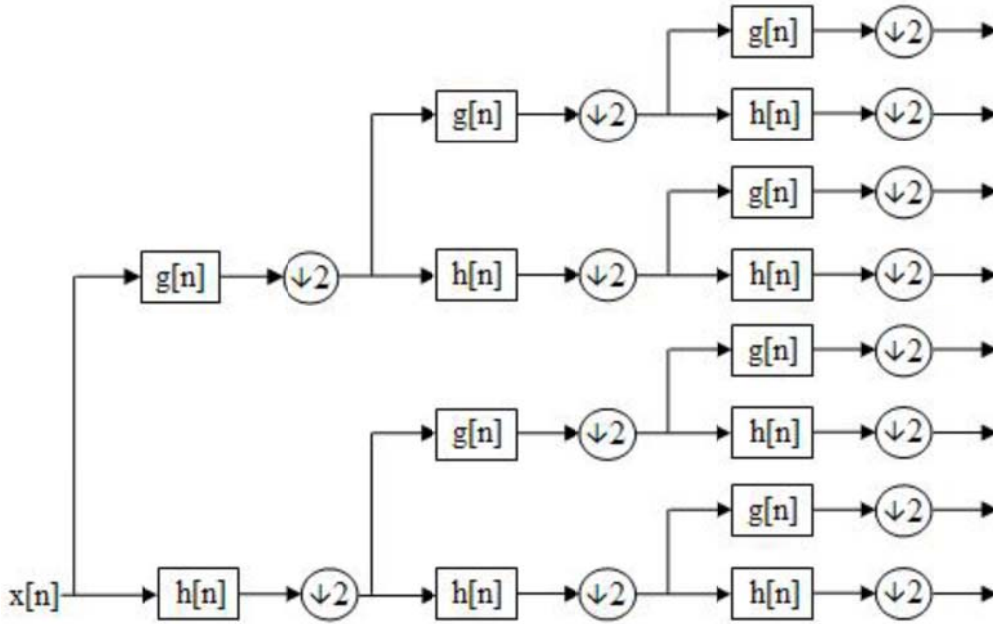
τ değeri ofset değerini, w açısal frekansı, t_j i. örnek zamanını, x ise yöntemin uygulandığı işareti ifade etmektedir. Bu yöntemde, τ ofset değeri ile $P_N(w)$ ifadesinin tüm t_j değerlerinin sabit bir değer ile kaydırılmasından bağımsız hale gelmesini sağlamaktadır (Clifford and Tarassenko 2005, Laguna et al. 1998). C ve MATLAB kodları PhysioNet internet sitesinden elde edilebilir (Goldberger et al. 2000).

2.7.2.3 Dalgacık Analizi Yöntemi

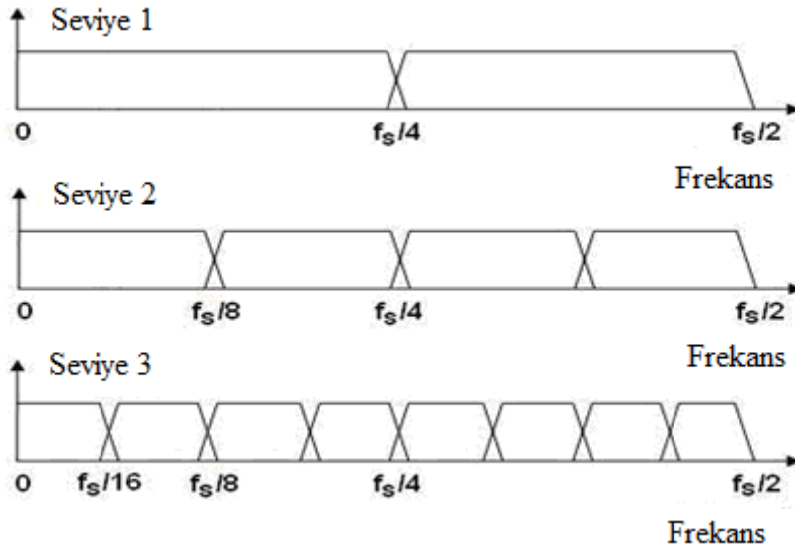
Fourier tabanlı yöntemlerde güç spektral yoğunluğunun hesaplanması için verilerin durağan olması gerekmektedir. Dalgacık analizi yönteminde ise verilerin durağan olması gerekmez. Fakat yeniden örnekleme gerek vardır.

Dalgacık paket dönüşümü ile yüksek ve alçak filtreler yardımıyla ana sinyal alt bileşenlerine detaylı olarak ayrılmıştır (Şekil 2.6). n seviyeli bir dalgacık paket dönüşümü ile 2^n farklı katsayı grubu elde edilir. Dalgacık paket dönüşümü sonucunda elde edilen frekans alanı

gösteriminde her bir seviyede frekans bantları yarıya düşmektedir (Şekil 2.7). Böylece işlem süresince örnek sayısının her seviyede yarıya düşürülmesiyle, işlem hızı artar ve gereksiz veriler atılmış olur (İşler 2009).



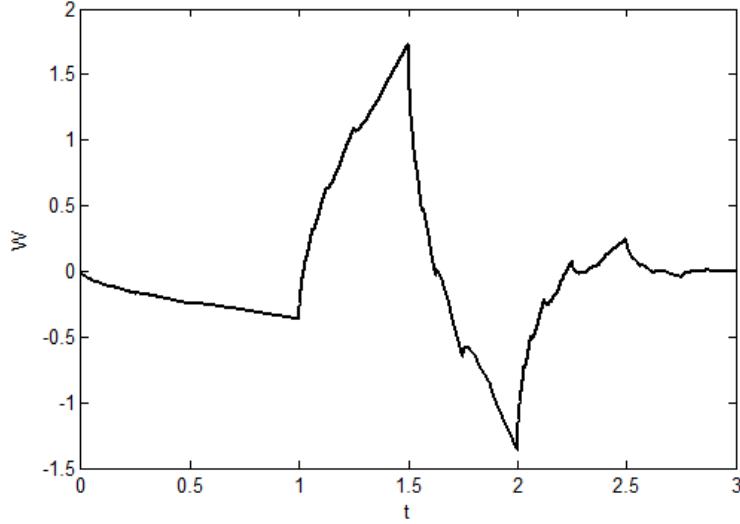
Şekil 2.6 Üç seviyeli dalgacık paket dönüşümü (İşler 2009).



Şekil 2.7 Dalgacık paket dönüşümünün frekans alanı gösterimi (İşler 2009).

Burada, f_s örnekleme frekansını göstermektedir.

Tüm dalgacık dönüşümü uygulamalarında en önemli aşama ana dalgacık fonksiyonunun seçimidir. Ana dalgacık olarak EKG ve KHD üzerine yapılan literatürdeki çalışmalarda EKG işaretine benzemesi sebebiyle çoğunlukla Daubechies 4 dalgacığı tercih edilmektedir (Şekil 2.8) (Asyalı 2003, Bakardjian and Yamamoto 1995, İşler ve Kuntalp 2007a, 2007b, 2010, Pichot et al. 1999, Shimazu et al. 1999, Torrence and Compo 1998, Wiklund et al. 1997).



Şekil 2.8 Daubechies 4 ana dalgacık fonksiyonu.

Dalgacık dönüşümü üzerinden dalgacık entropisi hesaplanması frekans alanı ölçümleri içerisinde en çok önerilen hesaplamadır (Rosso et al. 1998, Rosso et al. 2001; İşler ve Kuntalp 2007a, 2010). Bu hesaplamaları elde etmek için: i) dalgacık paket analizi katsayılarının elde edilmesi, ii) dalgacık enerjisinin hesaplanması ve iii) dalgacık entropisinin hesaplanması işlemlerinin aşamalı olarak yapılması gerekmektedir.

Dalgacık paket dönüşümü ile dalgacık paket analizi katsayıları elde edildikten sonra, her bir dalgacık katsayısının karesi alınarak buna karşılık gelen dalgacık enerjisi hesaplanır:

$$E_j = C_j^2 \quad (2.15)$$

burada C_j son dalgacık seviyesinin j -inci katsayısını göstermektedir (Rosso et al. 1998, 2001).

Daha sonra son seviyedeki dalgacık paketi katsayılarının sayısı N olmak üzere toplam dalgacık enerjisi:

$$E_{total}(m) = \sum_{j=1}^N E_j \quad (2.16)$$

şeklinde hesaplanır.

Normalize edilmiş dalgacık enerjisi değerleri

$$p_j = \frac{E_j}{E_{total}} \quad (2.17)$$

Formülüyle hesaplanır. Shannon entropi tanımından yola çıkarak (Shannon 1948) istenilen frekans bandı içinde kalan p_j değerleri ile:

$$WS_f = - \sum_{j \in f} (p_j \log_2(p_j)) \quad (2.18)$$

dalgacık entropisi değeri hesaplanır (Rosso et al. 2001).

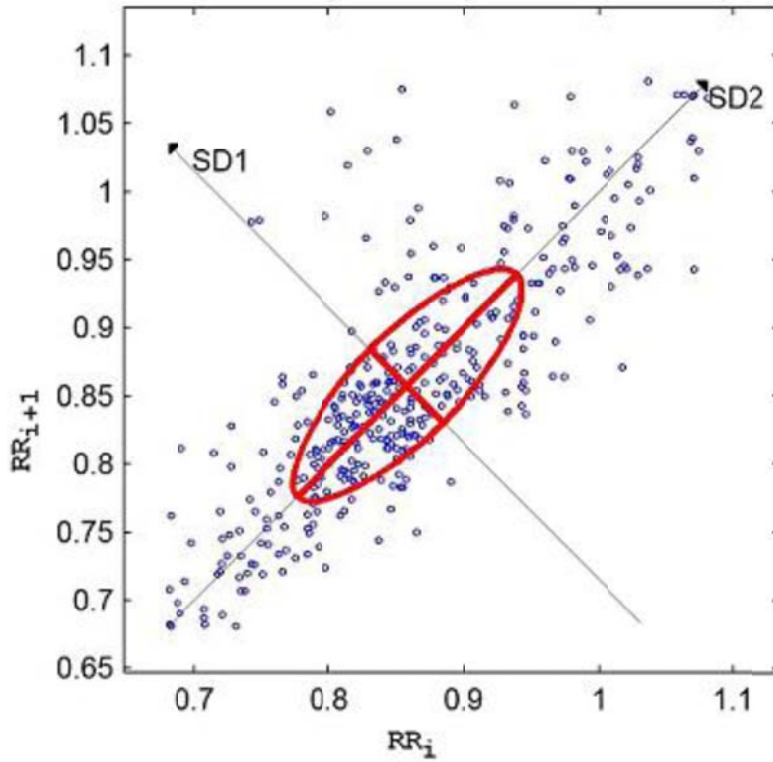
2.7.3 Doğrusal Olmayan Ölçümler

Doğrusal olmayan ölçümler, kalp vuruları (RR) arasındaki değişkenliğin çeşitli fizyolojik değişimlerden kaynaklanan değişimleri analiz etmeye yarayan ölçümlerdir. Bu şekliyle kalp hızı değişkenliği verileri doğrusal olmayan bileşenler içermektedir. Bu değerlerin bulunması ve dikkate alınması yapılan çalışmalar açısından önemlidir. Fakat henüz çok miktarda hasta üzerinde yürütülmüş ve klinik geçerliliği ispatlanmış bir yöntem geliştirilememiştir. KHD analizinde kullanılacak doğrusal olmayan yöntemler üzerine yapılacak çalışmalara ihtiyaç vardır (Task Force 1996).

Şu ana kadar literatürde varolan ölçümler ve KHD analizi üzerinde kullanılan doğrusal olmayan yöntemler arasında: Poincare çizimi (Kamen et al. 1996, Kamen and Tonkin 1995), Yaklaşık Entropi (ApEn) (Pincus 1991, Yeragani et al. 1998), Eğilim Yok Edilmiş Dalgalanma Analizi (DFA) (Hu et al. 2001, Kantelhardt et al. 2001, Peng et al. 1994, Taqqu et al. 1995), Örneklemeli Entropi (SampEn) (Goldberger et al. 2000), Çoklu Ölçekli Entropi Analizi (Costa et al. 2002, 2005) ve Sembolik Analiz (Mokikallio et al. 1997) vardır.

2.7.3.1 Poincare Çizimi

Doğrusal olmayan ölçümlerden poincare çizimi, herhangi işaretin yapısındaki doğrusal olmayan dinamiklerin belirlenmesinde kullanılan bir yöntemdir. Şekil 2.9 da gösterildiği gibi herbir verinin x eksenine ve kendisinden sonraki değeri y eksenine yerleştirildiğindeki çizimidir. Literatürde, kalp hızı değişkenliği analizi çalışmalarında kalple ilgili ayrıntılı bilgileri verebildiği ifade edilmiştir (Kamen et al. 1996, Kamen and Tonkin 1995, Woo et al. 1992). Poincare çiziminde elde edilen şekilsel ifade, veri ile ilgili dinamikleri göstermektedir (Berntson et al. 1997). Poincare çizimi ham veriyi kullanarak basit görsel gösterimi sağlaması ve özellikle kalp rahatsızlıklarını belirlemede faydalı olması nedeniyle sıkça tercih edilmektedir (Kamen 1996).



Şekil 2.9 SD1 ve SD2 ölçümleriyle birlikte örnek bir Poincare çizimi.

Poincare çiziminde elde edilen şekle elips uydurma yöntemi çok sık kullanılan yöntemlerden biridir (Marciano et al. 1994). Hesaplama nokta arasındaki mesafenin standart sapmaları ile elipsin genişliği (SD1) ve uzunluğu (SD2) vardır (Brennan et al. 2001). Ölçümlerin matematiksel ifadeleri (Landes et al. 1996):

$$SD_1 = \sqrt{\frac{1}{2}(SDSD)^2} \quad (2.19)$$

$$SD_2 = \sqrt{2(SD)^2 - \frac{1}{2}(SDSD)^2} \quad (2.20)$$

Şeklindedir. SD verilerin standart sapmasını ve $SDSD$ ise peş peşe gelen verilerin farkının standart sapmasını ifade etmektedir. Eğer peş peşe gelen veriler yerine gecikme değeri (lag)=1 kadar sonraki değerlere göre çizim elde edilmişse

$$SD_1 = std\left(\frac{x_{i+lag} - x_i}{\sqrt{2}}\right) \quad (2.21)$$

$$SD_2 = std\left(\frac{x_{i+lag} + x_i}{\sqrt{2}}\right) \quad (2.22)$$

Formülleri kullanılır.

Bununla birlikte, SD_1SD_2 çarpımı $\frac{SD_1}{SD_2}$ oranı (RATIO) bu bileşenler arasındaki ilişkiyi göstermektedir. Bu bileşenler otonom sinir sisteminin sempatik ve parasempatik aktivitelerinin dengesini ifade ettiği kabul edilmektedir (Kamen ve Tonkin 1995). Literatürde genellikle gecikme değeri (lag) 1 kabul edilmekte olmasına karşılık (Smith et al. 2007, Stein et al. 2008) bazı çalışmalarda bu değer 1 ile 10 arasında değişmektedir (Contreras et al. 2007, Thakre and Smith 2006).

2.7.3.2 Yaklaşık Entropi

Yaklaşık entropi lineer olmayan ölçümlere bir çözüm olması amacıyla Steve M. Pincus tarafından ortaya konulmuştur. Gürültülü ve kısa süreli veriler için oldukça başarılı sonuçlar vermektedir. Yaklaşık entropi ölçek değişmez, modelden bağımsız ve açık bir şekilde tespit edilmesi zor olan zaman serilerinin ayrımında kullanılmaktadır. Yaklaşık entropi, tepe oluşumlarını ya da genliklerdeki aralıklı davranışın altındaki değişimleri tespit eder (Pincus ve Viscarello 1992). Yaklaşık entropide veri içindeki karmaşıklık ya da düzensizliğe karşılık

gelen değerlerine, zaman serilerinde negatif olmayan bir değer belirler (Pincus 1991). Bu yöntemin en kullanışlı yönü kalp hızı işaretlerinde kullanılır olmasıdır. Bu sayede zaman serilerinin düzenlilik ve karmaşıklık miktarı ölçülür. Konuyla ilgili daha fazla bilgi ilgili referanslarda bulunmaktadır (Pincus 1991, Pincus and Viscarello 1992).

Yaklaşık entropi kalp hızı ölçümleri (Pincus et al. 1991) ve lineer olmayan zaman serileri, finans (Pincus and Kalman 2004), psikoloji (Pincus and Goldberger 1994) ve insan faktörleri mühendisliği alanlarında (McKinley et al. 2011) oldukça sık kullanılmaktadır. Ayrıca yaklaşık entropi EEG işaretlerinin sınıflandırılması, şizofreni (Sabeti 2009), epilepsi, (Yuan 2011) ve bağımlılık (Yun 2012) gibi hastalıkların teşhisinde de kullanılmaktadır.

2.7.3.3 Eğilim Yok Edilmiş Dalgalanma Analizi

Eğilim yok edilmiş dalgalanma analizi zaman serileri içindeki uzun mesafeli ilişkilerin ölçüsünde kullanışlı olduğu ortaya konmuştur (Kantelhardt et al. 2001). Öncelikle analizi yapılacak N tane örneğe sahip veri zaman serisine entegre edilir. Sonra entegre edilmiş zaman serileri eşit uzunlukta (n) parçalara bölünür. N uzunluktaki her bir parça, parçayı temsil eden eğime göre en küçük kareler hattına uygunlaştırılır. Bu hattın y eksenini $y_n(k)$ şeklinde ifade edilir. Bu entegre edilmiş ve eğimi yok edilmiş zaman serisinin RMS dalgalanma değeri aşağıdaki gibi hesaplanmaktadır.

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2} \quad (2.23)$$

Bu hesaplama ortalama dalgalanma $F(n)$ ve parça boyutu n arasındaki ilişkiyi tanımlamak için tüm zaman ölçekleri (kutu boyutları) üzerinde tekrarlanır. Genel olarak parça boyutu arttıkça hesaplama sonucu artacaktır. Logaritmik çizim üzerindeki lineer ilişki fraktal ölçekleme gücünün varlığını gösterir. Bu şartlar altında, dalgalanmalar $\log(n)$ ve $\log(F(n))$ ile ilişkili doğrunun eğimi üssel ölçekleme yapılarak tanımlanır.

2.7.3.4 Örnekleme Entropisi

Örnekleme entropisi deneysel ve klinik olarak elde edilmiş zaman seri verilerinin karmaşıklığını ve düzensizliğini ölçmek için kullanılan bir yöntemdir. Aslında örnekleme entropisi yaklaşık entropinin benzeri bir yapıya sahiptir. Ancak hesapsal olarak birbirinden farklıdır (Richman and Randall 2000). Örnekleme entropisi zaman serilerinin düzenliliğinin geliştirilmesine olanak sağlar. Yaklaşık entropi ile kıyaslandığında kardiyovasküler yapının lineer olmayan dinamiklerinin belirlenmesi için örnekleme entropisi oldukça kullanışlıdır.

2.7.3.5 Sembolik Analiz

Sembolik dinamikler, karmaşık sistemlerin ve ilgili sinyallerin dinamiklerinin belirlenmesini sağlayan analiz yöntemidir (Richman and Randall 2000). Gerçek gözlemlerin aralığı ya da gerçek gözlemlerin dönüştürülmüş aralığı sınırlı sayıda bölgelere ayrılır ve her bir bölge belirli bir sembolik değer ile ilişkilendirilir. Bu sembolik değer her bir gözlem ya da ardışık değerler arasındaki bölgelere göre belli sembollerle benzersiz bir şekilde eşleştirilir. Buradaki ilk eşleştirmeye statik dönüşüm denirken ikinci eşlemeye dinamik dönüşüm denir. Bu nedenle, gerçek gözlemler aynı uzunlukta bir dizi haline dönüştürülür. Dizideki elemanlar sadece birkaç farklı sembolden oluşmaktadır. Bu işleme sembolleştirme denmektedir. Genel olarak her bölümün ifade ettiği sembol tüm diğer semboller ile eş-muhtemel olmak zorundadır. Bu durum rastgele ve rastgele olmayan sembol dizileri arasında farklılıkları ortaya çıkarmak için yapılır.

BÖLÜM 3

YÖNTEM

Bu bölümde tez çalışmasının yürütülebilmesi için kullanılan veri tabanı ve yöntemler yer almaktadır. Sırasıyla, kullanılan veri tabanı, verilerden elde edilmiş öznitelikler, özniteliklerin seçimi, verilerin sınıflandırılması için kullanılan algılayıcı ağlar, çok katmanlı algılayıcı ağlar, destek vektör makinaları, radyal tabanlı fonksiyon ağlar, k en yakın komşu, doğrusal ayraç analizi, k ortalama kümeleme algoritması, genetik algoritma ile sınıflandırıcı eğitiminden ve son olarak sınıflandırıcı algoritmalarının başarımlarının değerlendirilmesi hakkında bilgi verilmiştir.

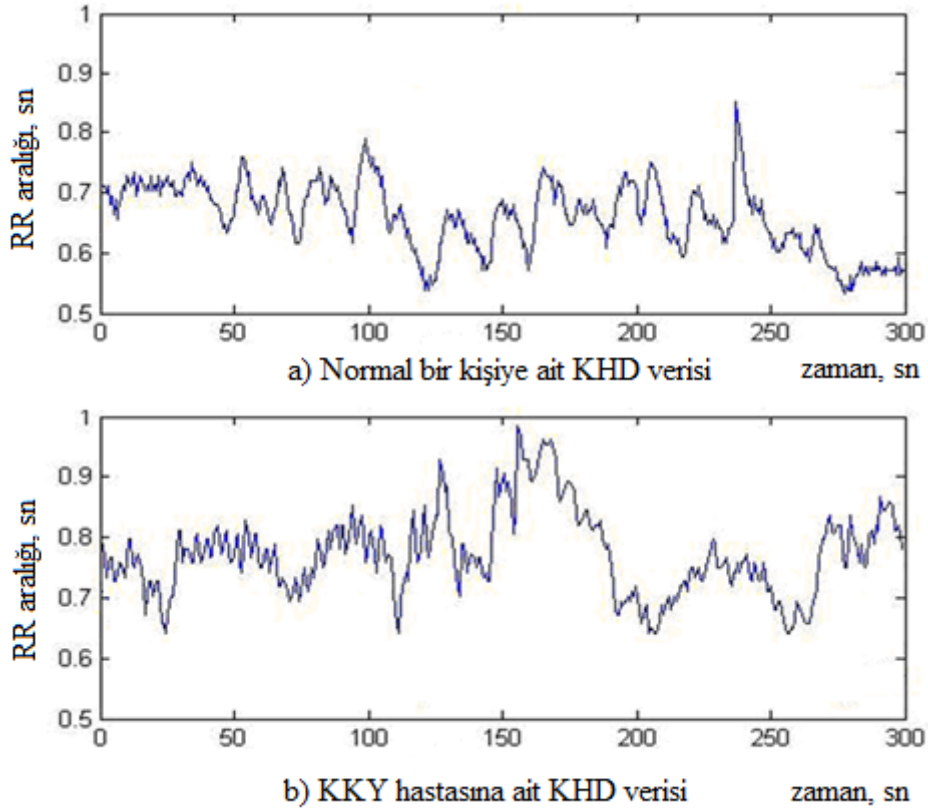
3.1 VERİLERİN ELDE EDİLMESİ

Yapılan çalışmada kullanılan KHD verileri MIT/BIH veri setleri <http://www.physionet.org> internet adresinden ücretsiz olarak elde edilen Physionet veri tabanından elde edilmiştir. Veri setinde normal kontrol grubu altında 54 sağlıklı bireyden ve 29 KKY hastası bireyden alınmış 24 saatlik Holter EKG kayıtları vardır. Üzerinde çalışılan bu veriler hakkında detaylı bilgi ilgili çalışmalarda vardır (Goldberger et al. 2000, Moody et al. 2001).

- **Konjestif Kalp Yetmezliği RR Aralıkları Veritabanı (CHF2DB):** Uzun süreli (24 saat) EKG kayıtlarından elde edilmiş, yaşları 34 – 79 arası değişen 8 tanesi erkek ve 2 tanesi kadın olup 19 hastanın cinsiyeti belli olmayan toplam 29 KKY hastasına ait verilerden oluşmaktadır. Elde edilen kayıtlar saniyede 128 örnek alınarak örneklenmiştir. Vuru bilgileri otomatik olarak elde edildikten sonra uzman hekimler tarafından tekrardan elden geçirilerek düzeltilmiştir.

• **Normal Sinüs Ritmi RR Aralıkları Veritabanı (NSR2DB):** Uzun süreli (24 saat) EKG kayıtlarından elde edilmiş, yaşları 28,5 – 76 arası değişen 30 tanesi erkek ve 24 tanesi kadın olan toplam 54 sağlıklı kişiden alınmış verilerden oluşmaktadır. Elde edilen kayıtlar saniyede 128 örnek alınarak örneklenmiştir. Vuru bilgileri otomatik olarak elde edildikten sonra uzman hekimler tarafından tekrardan elden geçirilerek düzeltilmiştir.

Her ne kadar veri tabanları 24 saatlik KHD verilerinden oluşsa da, sonuca daha hızlı ulaşabilmek ve klinik ortamda hastaları en az şekilde rahatsız etmek için sadece 5 dakikalık (300 saniyelik) KHD verileri kullanılmıştır. 300 saniyelik (a) normal kişiye ve (b) KKY hastasına ait iki adet KHD verisi gösterilmiştir (Şekil 3.1).



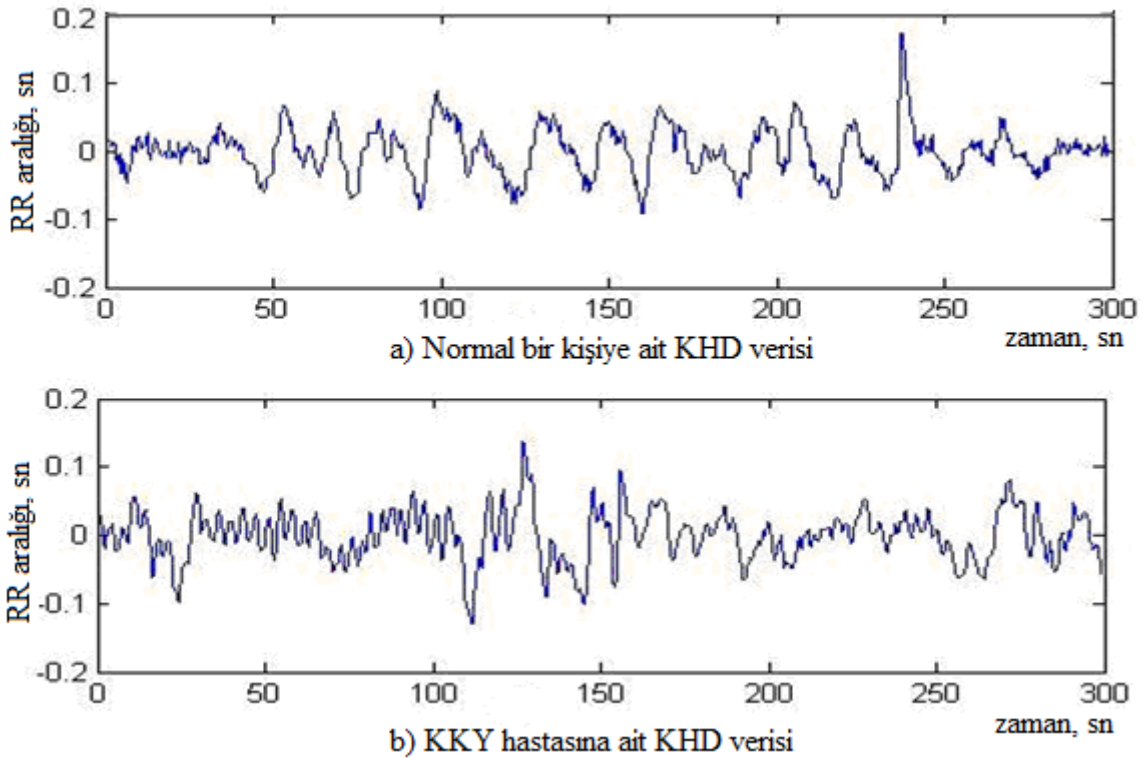
Şekil 3.1 KHD verileri.

3.2 ÖN İŞLEMLER

Ön işlem basamaklarında sırasıyla sadece ektopik vurunun yok edilmesi, yeniden örnekleme ve eğilim yok etme adımları uygulanmıştır. KHD analizinde kulakçık ve karıncık erken vuruları istenilmez, sadece normal vuruların kullanılması istenir. Bu nedenle istenilmeyen

vurular analizden çıkartılır (Clifford et al. 2006). Bu bilgi doğrultusunda Langley ve arkadaşları (Langley et al. 2001) uygulaması kolay bir algoritma geliştirerek erken vurular ve komşuları yok edilmiştir.

Çalışılacak veri içerisindeki istenmeyen vurular yok edildikten sonra, Bernston ve arkadaşlarının (1997) önerdiği 4 Hz örnekleme hızında (Clifford et al. 2006) tarafından önerildiği gibi kübik spline yöntemi ile veriler yeniden örneklenmiştir. Sonuç olarak eşit zaman aralıklarıyla örneklenmiş RR verileri elde edilmiştir. LS frekans alanı ölçümleri diğer frekans alanı ölçümlerinin aksine yeniden örnekleme ihtiyacı duymaz. Fourier ve dalgacık dönüşümü yöntemlerinde yeniden örnekleme gereklidir. Son olarak, RR verilerindeki eğilim yok edilerek veriler durağan hale getirilir. Durağan hale getirilmesi için $\lambda=1000$ parametresi ile Smoothness Priors yöntemi kullanılmıştır (Tarvainen, Ranta-aho ve Karjalainen 2002). Şekil 3.2’de ön işlem basamakları uygulanmış veri bulunmaktadır (İşler vd. 2005).



Şekil 3.2 Yeniden örneklenmiş ve eğilimi yok edilmiş KHD verisi.

3.3 ÖZİNİTELİKLERİN ÇIKARTILMASI

Doğrudan KHD verileri üzerinden yapılan ölçümler: Zaman alanı ölçümleri, Lomb-Scargle algoritması tabanlı frekans alanı ölçümleri, doğrusal olmayan ölçümlerden Poincare çizimi, Sembolik entropi, SampEn, ApEn ve DFA ölçümleri. Yeniden örneklenmiş KHD verileri üzerinden dalgacık entropisi, enerjisi ve varyansı ölçümleri yapılmaktadır. Yeniden örneklenmiş ve eğilimi yok edilmiş KHD verileri üzerinden FFT tabanlı frekans alanı ölçümleri yapılmaktadır. Ön işlem uygulanmış veya uygulanmamış veriler kullanılarak elde edilen ölçümler sonucundan 59 adet öznelik çıkartılmıştır. Bunlar, 1 adet hasta bilgisi, 4 adet zaman alanı ölçümü, 7 adet FFT tabanlı güç spektral yoğunluğu ölçümü, 23 dalgacık entropisi ölçümü, 7 Lomb-Scargle tabanlı güç spektral yoğunluğu ölçümü, 4 adet Poincare çizimi ölçümü, 10 adet sembolik entropi ile ApEn, SampEn ve DFA ölçümleridir.

Öznelikler çıkartıldıktan sonra tüm özneliklerin ortalama \pm standart sapma değerleri de çizelgelerde verilmiştir. Bununla birlikte çizelgelere eşit olmayan değişintili t-testi uygulanarak istatistiksel olarak karşılaştırılmış ve sonuçlar çizelgelerde P değeri olarak yazılmıştır. Çizelgelerdeki P değerine bakılarak eğer bu değer 0.05'ten daha küçük ise iki grup arasında istatistiksel olarak anlamlı bir fark olduğu sonucuna varılır. Eğer değer 0.05'ten büyükse iki grup arasında istatistiksel olarak anlamlı bir fark yoktur sonucuna varılır.

Çalışmada kullanılan veri setlerindeki hastaların yaşlarına ait bilgiler Çizelge 3.1'de verilmiştir.

Çizelge 3.1 Normal ve KKY hasta verilerine ait hasta bilgisi.

Öznelik	NORMAL	KKY	P
YAŞ	61.36 \pm 11.63	55.28 \pm 11.60	0.03

Zaman alanı ölçümleri içerisinde kullanılan MEAN, SDNN, RMSSD ve SDDSD ölçümleri Çizelge 3.2'de verilmiştir. İki grup arasındaki istatistiksel olarak anlamlı bir fark olup olmadığı P değeriyle görülmektedir.

Çizelge 3.2 Beş dakikalık veriler için elde edilen zaman alanı KHD analizi ölçümleri.

Öznitelik	NORMAL	KKY	P
MEAN	0.75±0.12	0.70±0.13	0.09
SDNN	0.05±0.02	0.03±0.03	0.00
RMSSD	0.02±0.01	0.02±0.01	0.01
SDSD	0.02±0.01	0.02±0.01	0.01

Frekans alanı ölçümlerinde FFT, Lomb-Scargle (LS) yöntemi ve dalgacık paket dönüşümü kullanılarak öznitelikler hesaplanmıştır. Dalgacık paket dönüşümü kullanılarak varyans (değişinti) (WV), enerji (WE) ve entropi (WS) hesaplamaları elde edilmiştir. Elde edilen bu ölçümler Çizelge 3.3'te verilmiştir. FFT ve Lomb-Scargle (LS) yöntemleri için eşit aralıklı 512 frekans noktası için değerler hesaplanmıştır. Frekans alanı ölçümleri arasında Lomb-Scargle tabanlı frekans alanı ölçümlerinin iki grubu (KKY ve NORMAL) ayırt etmede daha başarılı oldukları literatürdeki çalışmalarda rapor edilmiştir (İşler vd. 2008).

Çizelge 3.3 Kısa süreli (5 dakikalık) veriler için elde edilen frekans alanı ölçümleri.

Öznitelik	NORMAL	KKY	P
LS_TOTAL	688.4±102.2	692.3±110.3	0.88
LS_LFHF	5.96±3.82	2.58±2.46	0.00
LS_VLF	0.42±0.17	0.46±0.16	0.26
LS_LF	0.20±0.13	0.12±0.08	0.00
LS_HF	0.04±0.03	0.08±0.05	0.00
LS_NLF	0.81±0.11	0.60±0.21	0.00
LS_NHF	0.19±0.11	0.40±0.21	0.00
FFT_TOTAL	0.14±0.06	0.10±0.09	0.04
FFT_LFHF	6.74±2.86	6.08±4.31	0.40
FFT_VLF	0.00±0.00	0.00±0.00	0.00
FFT_LF	0.00±0.00	0.00±0.00	0.43
FFT_HF	0.00±0.00	0.00±0.00	0.60
FFT_NLF	0.84±0.07	0.81±0.08	0.04
FFT_NHF	0.15±0.07	0.19±0.08	0.04
WS_TOTAL	2.21±0.01	2.20±0.02	0.00
WS_LFHF	2.32±1.48	1.61±1.34	0.04
WS_ULF	2.19±0.01	2.19±0.01	0.05

Çizelge 3.3 Kısa süreli (5 dakikalık) veriler için elde edilen frekans alanı ölçümleri (devam ediyor).

Öznitelik	NORMAL	KKY	P
WS_VLF	0.01±0.01	0.01±0.01	0.00
WS_LF	0.01±0.01	0.00±0.01	0.00
WS_HF	0.01±0.00	0.00±0.00	0.01
WS_NLF	0.66±0.12	0.55±0.14	0.00
WS_NHF	0.34±0.12	0.44±0.14	0.00
WE_TOTAL	5188.5±1840.3	4876.6±2113.9	0.49
WE_LFHF	2.64±1.96	1.84±1.74	0.07
WE_ULF	5174.4±1832.0	4868.6±2099.3	0.49
WE_VLF	7.03±8.06	3.94±9.28	0.12
WE_LF	4.80±3.74	2.59±6.28	0.04
WE_HF	2.30±2.24	1.40±3.20	0.14
WE_NLF	0.33±0.13	0.44±0.16	0.00
WE_NHF	0.33±0.13	0.44±0.16	0.00
WV_LFHF	5.74±4.28	3.99±3.80	0.07
WV_ULF	2.24±2.57	1.34±2.46	0.13
WV_VLF	0.09±0.10	0.04±0.11	0.10
WV_LF	0.02±0.01	0.09±0.02	0.04
WV_HF	0.00±0.00	0.00±0.00	0.14
WV_NLF	0.81±0.09	0.72±0.12	0.00
WV_NHF	0.19±0.09	0.28±0.12	0.00

KHD analizi için doğrusal olmayan ölçümlerde çalışmaya dahil edilmiştir. Bu ölçümler arasında Poincare çizimi ölçümleri yer almaktadır. Bunlar, SD1, SD2 ve SD1/SD2 oranı (RATIO) gibi geleneksel Poincare çizimi ölçümlerinin ve SD1 ve SD2'nin çarpımı (SD1SD2) ölçümleri çalışmaya öznitelik olarak dâhil edilmiştir. Diğer doğrusal olmayan ölçümler arasında 1'den 10'a kadar farklı veri grupları ile ifade edilmesiyle oluşturulan sembolik entropi (SYM) ile SampEn ve ApEn gibi bazı entropi ölçümleri kullanılmıştır. Ayrıca DFA1 gibi frekansa göre güç spektral yoğunluğu arasındaki ilişkiyi veren grafiğin eğim bilgisi, doğrusal olmayan KHD ölçümü olarak çalışma eklenmiştir. Sonuç olarak tüm bu yöntemlerle elde edilen doğrusal olmayan KHD ölçümlerine ait değerler Çizelge 3.4'de verilmiştir.

Çizelge 3.4 Kısa süreli (5 dakikalık) veriler için doğrusal olmayan yöntemlerle elde edilen KHD ölçümleri.

Öznitelik	NORMAL	KKY	P
SD1	0.02±0.01	0.01±0.01	0.01
SD2	0.07±0.03	0.04±0.04	0.00
SD1SD2	0.00±0.00	0.00±0.00	0.13
RATIO	0.27±0.11	0.40±0.18	0.00
SYM01	2.00±0.42	2.00±0.42	0.99
SYM02	2.44±0.46	2.37±0.50	0.55
SYM03	2.62±0.42	2.78±0.41	0.10
SYM04	2.80±0.44	2.94±0.47	0.17
SYM05	2.88±0.43	3.04±0.49	0.13
SYM06	2.91±0.39	3.13±0.48	0.03
SYM07	2.91±0.37	3.16±0.42	0.01
SYM08	2.86±0.35	3.12±0.49	0.01
SYM09	2.79±0.32	3.07±0.52	0.00
SYM10	2.73±0.32	2.94±0.47	0.02
DFA1	1.02±0.15	1.05±0.19	0.42
SampEn	1.28±0.41	1.47±0.42	0.06
ApEn	0.01±0.01	0.00±0.02	0.56

3.3.1 Özniteliklerin Ölçeklenmesi

Sınıflandırma ve kümeleme algoritmalarında kullanılan özniteliklerin birimlerinin ve birbirlerine göre büyüklüklerinin çok farklılıklar göstermesi nedeniyle, kullanılan özniteliklerin aynı ölçeğe indirgenmesi gerekmektedir. Bu işleme normalizasyon adı verilir (Duda et al. 2001). Literatürde yapılan çalışmalarda özniteliklerin ölçeklenmesinin sonuçları etkilediği görülmektedir (İşler ve Kuntalp 2007a, 2007b, 2010). Bu çalışmada [0,1] aralığına ölçekleme olan (Min-Max) yöntemi kullanılmıştır:

$$X_k^{yeni} = \frac{X_k - \text{Min}\{X_k\}}{\text{Max}\{X_k\} - \text{Min}\{X_k\}} \quad (3.1)$$

Burada, X_k k-inci öznitelik olmak üzere $\text{Min}\{X_k\}$ ve $\text{Max}\{X_k\}$ sırasıyla bu özneliğe ait en küçük ve en büyük değerleri göstermektedir.

3.4 ÖZNETELİK SEÇİMİ

Öznitelik sayılarının artması ek ölçümleri gerektirdiğinden zaman alıcı ve masraflı bir iştir. Bu nedenle toplam değişintiyi en az sayıda değişkenle ve örnekle açıklamak esas amaçtır (Efe vd. 2000).

25 yıl öncesine kadar 40'tan fazla öznitelik (değişken) kullanan çalışma sayısı oldukça azdı dolayısıyla değişken ya da örnek seçimi ile ilgili az sayıda çalışma bulunmaktaydı (Guyon and Elisseeff 2003). Fakat günümüzde bu durum çok değişti. Veri miktarlarının, ölçüm çeşitliliğinin ve dolayısıyla öznitelik sayılarının artmasıyla sınıflandırıcıların tasarımları zorlaştı ve sınıflandırıcı başarımları olumsuz etkilendi (Duda et al. 2001). Bu problemi çözmek için öznitelik / örnek seçimi yöntemlerine ihtiyaç duyuldu. Literatürde kullanılan birçok yöntem bulunmaktadır.

Öznitelik seçiminde sıkça kullanılan yöntemler arasında olan: Temel bileşen analizi, geriye doğru eleme yöntemi ve bağımsız örneklem t-testi yöntemleri bu çalışmamızda kullanılmıştır. Her araştırmacının kolayca uygulayabilir olması, hesaplama ağırlığının az olması ve en önemlisi hızlı sonuç vermesi bu yöntemlerin tercih edilmesine sebep olmuştur. Sonraki başlıklarda kullanılan bu yöntemler açıklanacak ve bu yöntemler sonucunda elde edilen veriler kullanılarak başarımların sonuçları analiz edilecektir.

3.4.1 Temel Bileşen Analizi

Temel Bileşenler Analizi (TBA) orijinal verilerin değişinti (varyans) yapısını daha az sayıda ve bu değişkenlerin doğrusal bileşenleri olan yeni değişkenlerle ifade etme yöntemidir (Shlens 2009). Toplam değişintinin yeni değişkenlerde hangi oranda göz ardı edileceği önceden belirlenir. Bu çalışmada bu oranın etkisi araştırılacağı için, çalışmada %0 (TBA yok), %1, %5, %10 ve %20 olarak bu oran seçilmiştir. Burada %1 ifadesi, değişintideki artış %1'in altına düşüncüye kadar temel bileşen araştırmasına devam edileceği anlamına gelmektedir (Duda et al. 2001). Böylece özniteliklerin sayısı (boyutu) azaltılmış olur.

3.4.2 Geriye Doğru Eleme Yöntemi

Öncelikle tüm özniteliklerin dâhil edildiği sınıflandırıcı başarımı tespit edilir. Daha sonra her adımda bir öznitelik listeden çıkarılarak sınıflandırıcı başarımına olan etkileri bulunur. Sınıflandırıcı başarımını hiç düşürmeyen veya en az düşüren öznitelik tespit edilerek listeden kalıcı olarak çıkarılır. Bu işlem elenmesi için seçilen tüm özniteliklerin sınıflandırıcı başarımına olan katkılarının önemli olması durumunda sona erdirilir.

3.4.3 Bağımsız Örneklem t-Testi

Bağımsız Örneklem t-Testi iki ortalamanın karşılaştırılması için kullanılan bir analiz yöntemidir. Bu analizin yapılabilmesi için öncelikle analizi yapılacak sınıflara ya da gruplara ait verilerin elde edilmesi gerekmektedir. Elde edilen bu veriler arasındaki anlamlı farklılıkların olup olmadığını test etmek için p değerlerine bakılmaktadır. P değerinin bulunması için IBM SPSS 19 programı kullanılmıştır. P değerini belirlemeden önce iki sınıf arasında bir hipotez belirlemek gerekmektedir. Örneğin, “normal bireylerle KKY hastası bireyler arasında cinslerin önemi yoktur.” şeklinde hipotez belirlenir. Bu hipoteze göre elde edilen p değerine göre hipotezin doğruluğu araştırılır. Buna H_0 hipotezi denir. Bu hipotez sınıflar arası farksızlığı esas alırken kurulan bu hipotezin aksine kurulan H_1 hipotezi de sınıflar arası fark olduğunu iddia etmektedir.

$$H_0 \Rightarrow \mu_1 = \mu_2 \text{ (veya } \mu_1 - \mu_2 = 0 \text{)} \quad (3.2)$$

ve

$$H_1 \Rightarrow \mu_1 \neq \mu_2 \text{ (veya } \mu_1 - \mu_2 \neq 0 \text{)} \quad (3.3)$$

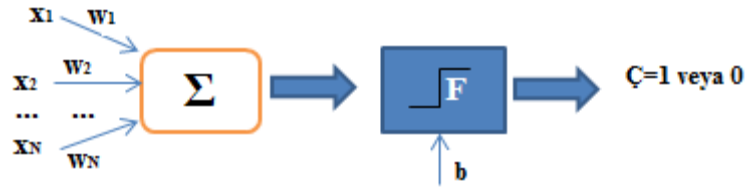
Burada μ_1 ve μ_2 değerleri sınıfların ortalamalarını göstermektedir. Hipotezin doğruluğuna bakmak için ihtiyaç duyulan önemli parametrelerden biri “anlamlılık düzeyi” dir. Bu değer analizi yaparken kullanıcıya bağlıdır. %’de kaçlık yanılma ihtimali göz ardı edilebilir. Kurulan H_0 hipotezinin doğruluğunu test etmek için p değeri 0,05 alınır, elde edilen sonuç belirlenen anlamlılık değerinden büyükse H_0 hipotezi kabul edilir aksi halde H_0 hipotezi reddedilir H_1 hipotezi kabul edilir.

Yapılan bu çalışmada 0.01 (%1), 0.02 (%2), 0.05 (%5), 0.1 (%10) ve 0.2 (%20) anlamlılık değeri için çalışma gerçekleştirilmiştir. Bu değerlere göre öznelik seçimi yapılmıştır.

3.5 SINIFLANDIRICILAR

3.5.1 Algılayıcı Ağlar

Sınıflandırma amacıyla geliştirilmiş, bir sinir hücresinin birden çok giriş için çıkış üretmesini temeline dayanan bir yapıya sahiptir. Ağın çıktı değeri, ağırlıkla çarpılarak elde edilen giriş değerlerinin eşik değeriyle toplanması sonucu elde edilir. Basit bir algılayıcı yapısı şekilde görülmektedir (Öztemel 2003) (Şekil 3.3).



Şekil 3.3 Algılayıcı yapısı (Öztemel 2003).

Şekilde de görülmektedir ki algılayıcı tek bir yapay sinir hücresinden oluşmaktadır. Bu sinir hücresinin eğitimi w (ağırlık) değerlerinin değişmesiyle elde edilmektedir. Yani her bir girdi değerine karşı ağa çıkış değerleri gösterilip daha sonra öğrenme kuralına göre ağın çıktı değerleri hesaplanır. Şayet ağın çıktı değeri olması gerekenden farklı ise ağın ağırlık değeri (w) ve eşik değeri (b) değiştirilir eğer hesaplanan çıktı ile beklenen çıktı aynı ise ağırlıklarda ve eşik değerde herhangi bir değişiklik yapılmaz. Bu şekilde ağ eğitilmiş olur (Rosenblatt 1957, Liou et al. 2013).

Bir algılayıcının çıkış fonksiyonu şu şekilde hesaplanır.

$$\text{Çıkış} = \begin{cases} 1 & \text{eğer } w \cdot x + b > 0 \\ 0 & \text{eğer } w \cdot x + b < 0 \end{cases} \quad (3.4)$$

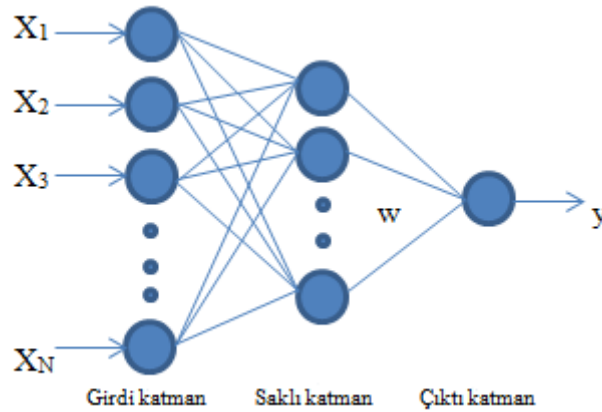
Burada x girdi verisi, w ağırlık değeri ve b ise eşik değeridir. Bu değerler kullanılarak çıkış değeri bulunmaktadır.

Eđitilen bu ađ daha sonra dıřardan yeni girilen bir veriyi, ađın eđitimi sonucunda elde edilen ađırlık deđerlerini kullanarak hangi sınıfa ait olduđunu tespit edebilir. Bu sayede sınıflandırma iřlemi yapılmıř olur.

Bu alıřmada yapı olarak basit olması ve hesaplama yknn fazla olmaması nedeniyle ok katlı sınıflandırıcı tasarımlarının ilk iki basamađındaki n sınıflandırıcıların tasarımında tercih edilmiřtir.

3.5.2 ok Katmanlı Algılayıcı Ađlar

Yapay sinir ađı uygulamalarında ok kullanılan ađ yapısıdır (Duda et al. 2001). Hem dođrusal uygulamalarda hem de dođrusal olmayan uygulamalar iin kullanılabilir. MLP giriř katmanı, bir veya birden ok saklı katman ve ıkıř katmanından oluřmaktadır (řekil 3.4).



řekil 3.4 ok katmanlı algılayıcıların yapısı.

MLP'nin her bir katmanında birok sinir hcreleri bulunmaktadır. Bu sinir hcreleri kendilerinden nceki ve sonraki katmanlardaki sinir hcrelerine belirli ađırlıklarla (arpan katsayılarıyla, w) bađlıdır. Bu ađırlıklarında kullanılmasıyla girdi verisine karřılık retilmesi gereken ıktı deđerinin ađa gsterilmesi ve bunun sonucundan her girdi verisine karřı ađın dođru ıktı deđerini retebilmesi MLP'nin eđitimi olarak ifade edilir. Yani, bu ađırlık deđerlerinin kullanılmasıyla beklenen ıktı ile ađın ıktısı arasındaki hata minimum olana kadar ađın eđitimi gerekleřtirilir. Bu đrenme kuralına Genelleřtirilmiř Delta đrenme Kuralı denilmektedir (ztemel 2003). Genel olarak ađın ıktı fonksiyonu ve hata fonksiyonu:

$$y_{\zeta_i} = f\left(\sum_{i=1}^N w_{ji} x_i\right) \quad (3.5)$$

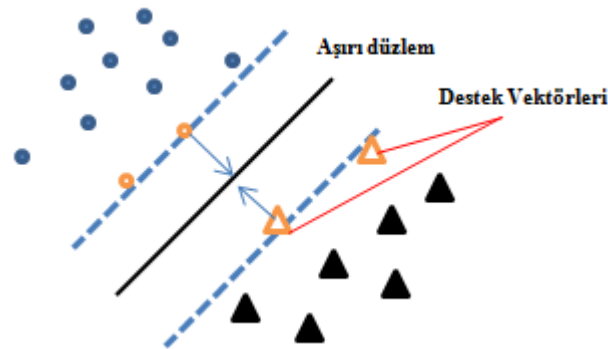
$$E = \frac{1}{2} \sum_i (y_{b\zeta_i} - y_{\zeta_i})^2 \quad (3.6)$$

Burada x_i girdi verisi, w_{ji} ağırlık, $f(\cdot)$ aktivasyon fonksiyonu, y_{ζ_i} i. ağın çıktısı, $y_{b\zeta_i}$ i. beklenen çıktıdır.

MLP ağına eğitiminde Levenberg-Marquardt (LM) yöntemi kullanılmıştır. LM, en az kareler yöntemi üzerine kurulmuş bir optimizasyon yöntemidir.

3.5.3 Destek Vektör Makinaları

Destek vektör makinaları (SVM) istatistiksel öğrenme teorisine dayalı bir yöntemdir (Vapnik 1998). SVM sınıflandırma ve regresyon amaçlı hem doğrusal hem de doğrusal olmayan veri gruplarında kullanılmaktadır (Byun and Lee 2003). Bu yöntemde amaç iki grubu bir birinden ayıran düzlemin bulunmasıdır. Bunun için iki gruba da yakın ve birbirine paralel iki sınır çizgisi belirlenir. Bu sınır çizgilerine destek vektörleri denmektedir. Bu sınıflandırıcı iki sınır çizgisinden yola çıkarak ortak bir düzlemin belirlenmesi işlemine dayanmaktadır (Şekil 3.5). SVM’de bazı veriler doğrusal bir şekilde ayıramamaktadır. Bu durumda ise çekirdek fonksiyonları kullanılarak veriler başka bir boyuta taşınıp doğrusal olmayan dönüşümler de kullanılabilir.

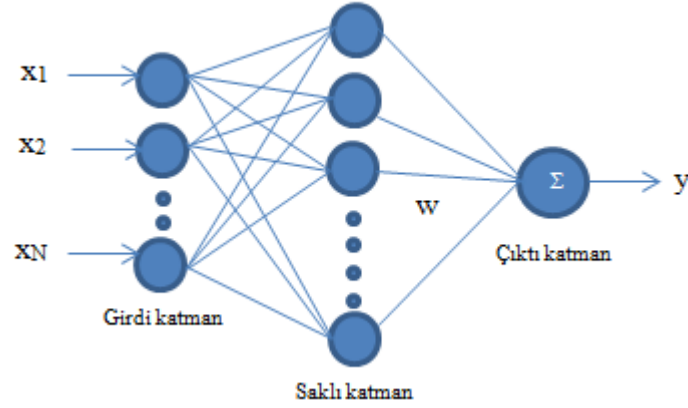


Şekil 3.5 SVM destek vektörleri ve aşırı düzlem.

3.5.4 Radyal Tabanlı Fonksiyon Ağlar

Radyal Tabanlı Fonksiyon (RBF) yapay sinir ağları da girdi katmanı, saklı katman ve çıktı katmanı olmak üzere üç katmandan oluşmaktadır (Şekil 3.6). MLP ağlarına oranla daha hızlı çalışırlar (Duda et al. 2001). RBF diğer bir ifadeyle yapay sinir ağlarının çok boyutlu bir uzayda eğri uydurma yaklaşımıdır. Bu ağ türünün saklı katmanında aktivasyon fonksiyonu olarak Gauss fonksiyonu kullanılmaktadır. Gauss aktivasyon fonksiyonu ifadesi:

$$y_i(x) = \sum_{j=1}^k w_{ji} \exp\left(-\frac{\|x - \mu_j\|^2}{2\sigma_j^2}\right) \quad (3.7)$$



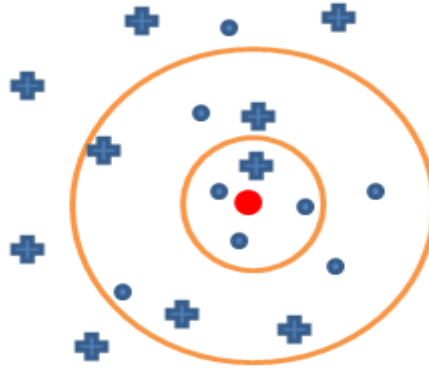
Şekil 3.6 RBF yapay sinir ağının yapısı

Burada w_{ji} ağırlıkları, k gizli katmandaki nöron sayısı, x girdi verisi, μ_j j 'inci nöron merkezi, σ_j j . dağılım parametresi olmak üzere y_i ağın çıktısını vermektedir. Ağın çıktısı saklı katman çıkışlarının w_{ji} ağırlıkla çarpımı sonucu elde edilir ve amaç optimum ağırlık değerlerinin bulunmasıdır.

3.5.5 K En Yakın Komşu

Literatürde geliştirilmiş birçok sınıflandırıcı algoritması bulunmasına rağmen, bu çalışmada basitliği ve kolay anlaşılabilirliği nedeniyle k en yakın komşu (KNN) sınıflandırıcısı kullanılmıştır. KNN örnek tabanlı bir sınıflandırıcı olup sınıflandırıcı sistem parametreleri olarak sadece örnekler tutulmaktadır. Bu algoritma tüm örneklerin d boyutlu uzaydaki noktalar olduğunu kabul eder (Mitchell 1997). Sınıfı bilinmeyen yeni bir örnek test

edildiğinde, test edilen örneğin kendi komşularından k tanesinin çoğunluğunun bulunduğu sınıfa atanır (Şekil 3.7). Bu şekilde kırmızı ile gösterilen örneğin + mi yoksa O grubundan mı olduğu bulunmak istenmektedir. Eğer $k = 4$ seçilirse, 3 tane O olduğu için test örneğinin O sınıfından olduğu, aynı şekilde eğer $k = 12$ seçilirse 7 tane O olduğu için test örneğinin O sınıfından olduğu bulunacaktır (İşler 2009). k genellikle küçük değerli bir pozitif tam sayıdır. İki sınıflı problemlerde, k değerinin tek sayı olması, test edilen örneğin sınıfını belirlemede kararlı bir yapı oluşturur (Nixon and Aguado 2002). Aksi takdirde test edilen örnek her iki sınıfa eşit sayıda yakın komşuya sahip olabilir.



Şekil 3.7 KNN sınıflandırıcı örneği.

Bu sınıflandırıcı algoritmasında önceden sınıfı bilinen örneklerle kıyaslama yapıldığından, ayrıca bir eğitim aşaması yoktur. Diğer bir deyişle, sınıflandırıcının eğitim aşaması sadece örneklerin ve ait oldukları sınıf bilgilerinin saklanmasıdır. Komşuların belirlenmesinde, örneklerin hepsi çok boyutlu uzaydaki vektörler olarak gösterilir. Genellikle Öklit mesafesi ölçütü kullanılsa da, Mahalanobis mesafesi gibi diğer ölçütler de kullanılabilir (Vatanabe 1985).

3.5.6 Doğrusal Ayraç Analizi

Doğrusal ayraç analizi (LDA) iki ya da daha fazla gruba ait özneliklere göre ayırım yapabilen istatistiksel bir analiz yöntemidir. Matematiksel olarak ayraç fonksiyonları kullanılmaktadır. LDA sınıflandırma ve boyut azaltma yöntemlerinde oldukça yaygın kullanıma sahiptir (Fukunaga 1990).

3.5.7 K Ortalama Kümelemesi

K ortalama algoritması adından da anlaşılacağı gibi giriş uzayını K adet merkezle ifade etmeye çalışan bir yöntemdir (Duda et al. 2001, Martis and Chakraborty 2011, et al. 2011). Merkezlere ilk değer ataması rasgele olarak yapıldıktan sonra merkez değerlerinin güncellenmesi için iki farklı yöntem kullanılır. Birinci yöntemde (batch metodu) giriş kümesindeki her bir örneğin hangi merkeze yakın olduğu hesaplanır. Aynı merkeze yakın olan örneklerin ortalaması alınarak merkezin değeri güncellenmiş olur. Durma koşulu sağlanana kadar bu işlem tekrar edilir. İkinci yöntemde (online metodu) giriş kümesinden bir örnek seçilir ve bu örneğin merkezlere olan uzaklığına bakılır. Örneğin en yakın olduğu merkez bulunarak bu merkezin değeri güncellenir. Her bir örnek için bu işlem tekrarlanır. Merkez değeri güncellenirken merkezle örnek arasındaki mesafe değeri her adımda azalan bir öğrenme katsayısıyla çarpılarak kullanılır. Bu sayede ilk adımlarda merkezlerin yer değiştirmesi büyük miktarlarda olurken zamanla yer değiştirme azalır ve merkezler yakınsar. Durma şartı sağlanıncaya kadar her örnek için bu işlem tekrarlanır. Örneklerin işleme sırası her adımda aynı olacağı gibi rasgele sırada da olabilir (Duda et al. 2001). Bu çalışmada birinci yaklaşım tercih edilmiştir.

3.6 GENETİK ALGORİTMA İLE SINIFLANDIRICI EĞİTİMİ

Genetik Algoritma (GA) kâinattaki evrim prensiplerini kullanarak geliştirilen bir en iyileme yöntemidir. Bireylerin hayatta kalabilmeleri ve neslini devam ettirebilmeleri, onların içinde buldukları çevreye uyum sağlamaları ile mümkündür. GA yapay olarak toplum genetiği prensiplerini kullanan, verilen bir problemin çözümlerini rasgele yönlendirilmiş ve paralel arama yöntemiyle bulan evrimsel algoritmadır (Goldberg 1989, Holland 1975). GA'nın en büyük avantajlarından biri arama uzayının çok büyük olduğu, doğrusal olarak değişmediği, gürültülü ve çok iyi tanımlanmamış olan problemlerde bile çözüm üretebilir olmasıdır (Koza 1992).

GA, problemlere çözüm üretmek, modelleme ve eniyileme uygulamaları için kullanılmaktadır. Örnek olarak optimizasyon, mühendislik tasarımları, planlama, yapay zekâ uygulamaları, elektronik devre tasarımı, makine öğrenmesi, robot öğrenmesi ve örüntü tanıma ve daha birçok kullanım alanı vardır. GA, diğer eniyileme yöntemleri arasında oldukça etkili ve problemlere tek bir çözüm üretmek yerine farklı çözümlerden oluşan bir çözüm kümesi

üretmesi ve aynı zamanda hiçbir çözüm tekniği bulunmayan problemlere çözüm araması nedeniyle günümüzde oldukça sık kullanılmaktadır. GA yöntemini diğer geleneksel arama ve en iyileme yöntemlerinden ayıran en temel etkenler şu şekilde sıralanabilir (Koza 1992):

- GA tasarım parametrelerinin (değişkenlerin) değerleriyle değil, onların kodlanmış halleriyle çalışır, böylece parametreler kodlanabildiği sürece çözüm üretilebilir.
- GA'nın yerel minimumlara takılması engellenmek için tek bir çözüm ile değil bir çözüm kümesi ile çalışır.
- GA uygulamaları problemden bağımsız hale gelmesi için GA bir uygunluk fonksiyonu ile çalışmakta olup fonksiyonun türevlerine ihtiyaç duymazlar.
- GA tahmini veya olasılığa dayalı olmayan değil rasgele arama yöntemidir.

GA yukarıda saydığımız uygulama alanları dışında literatürde daha çok en iyi öznelik kümesini seçmek ve sınıflandırıcı parametrelerinin belirlenmesinde kullanılmaktadır (Raymer et al. 2000). Bu çalışmada ise GA, çok katlı sınıflandırıcı tasarımlarında ön sınıflandırıcı olarak kullanılan algılayıcı ağların en iyi ağırlık değerlerinin bulunmasında kullanılmıştır. Bulunan bu ağırlık değerleri ile sınıflandırma doğruluğu daha yüksek algılayıcı ağlar elde edilmiştir. Bulunan en iyi ağırlık değerine göre ağ test edilecek ve sınıflandırma işlemi tamamlanmıştır.

3.7 SINIFLANDIRICI BAŞARIMLARININ BELİRLENMESİ

Amerikan Kalp Derneği hastalık teşhisinde dört temel ölçüm önermektedir (Gibbons et al. 1997). Bu ölçümler benzer diğer çalışmalarda da kullanılması sebebiyle teşhis amaçlı sınıflandırıcılarda da kullanılmaktadırlar (Valafar 2001). Ölçüm yaparken, sınıflandırıcı ile test edilen veri için negatif ya da pozitif kararlarından biri verilir. Verilen bu kararın sonunda da doğru ve yanlış olmak üzere iki tür sonuç üretilir. Böylece iki sınıfa ait veriler için, dört farklı olası durum ortaya çıkmaktadır. Örnek pozitif ise ve pozitif olarak sınıflandırılmışsa doğru pozitif (TP), negatif olarak sınıflandırılmışsa yanlış negatif (FN) olmaktadır. Benzer şekilde eğer örnek negatif ise ve negatif olarak sınıflandırılmışsa doğru negatif (TN) ve pozitif olarak sınıflandırılmışsa (FP) olmaktadır. Çalışmaya göre bu ifadeler yazılırsa; gerçekten hasta olup sınıflandırıcı tarafından hasta olarak tespit edilenlere TP ve yanlışlıkla sağlam olarak tespit edilenlere FN, gerçekten sağlam olup sınıflandırıcı tarafından sağlam

olarak tespit edilenlere TN ve yanlışlıkla hasta olarak tespit edilenlere FP denir. Bu durumlar için 2x2'lik bir karar matrisi oluşturulabilir (Eberhart and Dobbins 1990). Şekil 3.8 matrisin gösterimi vardır. Bu değerler kullanılarak sınıflandırıcı performansları için ACC (genel başarımlık), SEN (hassaslık), SPE (özgünlük), POS (pozitif değerililik) ve NEG (negatif değerililik) değeri hesaplanmıştır.

		Sınıflandırıcı Kararı	
		Normal	KKY
Gerçek durum	Normal	TN	FP
	KKY	FN	TP

Şekil 3.8 İkiye ikilik karar matrisi.

Her türlü sınıflandırıcı problemi için bu değeri hesaplandıktan sonra sınıflandırıcı performansları hesaplanabilir (Eberhart and Dobbins 1990, Duda et al. 2001):

$$ACC = doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.8)$$

$$SEN = hassaslık = \frac{TP}{TP + FN} \quad (3.9)$$

$$SPE = özgünlük = \frac{TN}{TN + FP} \quad (3.10)$$

$$POS = pozitif seçicilik = \frac{TP}{TP + FP} \quad (3.11)$$

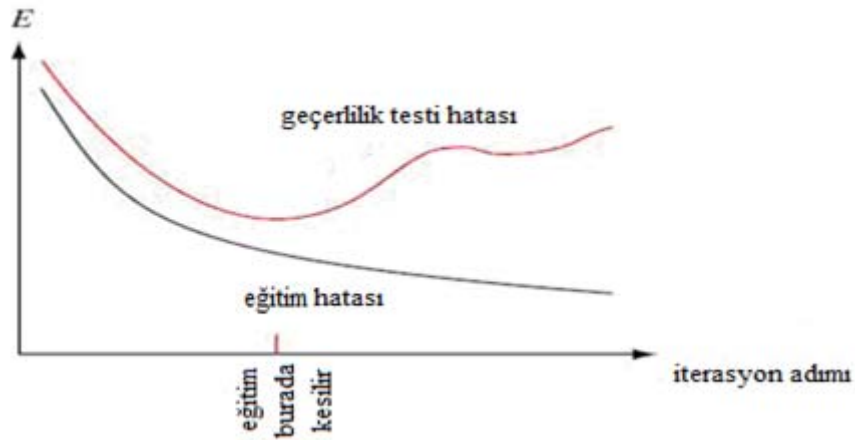
$$NEG = negatif seçicilik = \frac{TN}{TN + FN} \quad (3.12)$$

3.7.1 Çapraz Geçerlilik

Çapraz geçerlilikte üzerinde çalışılacak veri seti iki gruba ayrılır. Bunlardan birincisi (eğitim seti) sınıflandırıcının model parametrelerinin belirlenmesinde kullanılırken diğeri (test seti) ise eğitilen sınıflandırıcının başarımlık performansını ölçmek yani genelleme hatasının

tahmininde kullanılır. Sonuçta hedef istenen düşük genelleme hatasına erişmek olduğundan, sınıflandırıcının eğitimi en düşük geçerlilik hatasına ulaşıncaya kadar sürdürülür (Şekil 3.9). Bununla birlikte eğitimde kullanılan verilerin test kümesinde kullanılmaması başarımların hesaplanmasında oldukça önemlidir (Duda et al. 2001).

Yukarıda açıklanan temel genelleme yöntemine *m-parçalı çapraz-geçerlilik* (m-fold cross-validation) denilmektedir. Bu durumda toplam n örneğin bulunduğu eğitim veri seti, her birinde $\frac{n}{m}$ örneğin bulunduğu m ayrı parçaya ayrılır. Her seferinde farklı bir veri seti kümesi geçerlilik için ayrılacak şekilde sınıflandırıcı m defa eğitilir. Böylece elde edilen m adet hatanın ortalaması ile sınıflandırıcı performansı tahmin edilmiş olur. Eğer $m=n$ olarak alınırsa, yöntem *birisi-dışarıda* (leave-one-out) yaklaşımına dönüşür. Bu yaklaşımda, sınıflandırıcı her seferinde sadece bir örnek hariç olacak şekilde kalan tüm veri seti ile eğitilir ve bu işlem tüm örnekler için tekrarlanır. Yüksek miktarda örneğin yer aldığı durumlarda hesaplama yükü aşırı olmaktadır. Bu nedenle, özellikle veri setinin küçük olduğu durumlarda bu yöntemle çapraz geçerlilik testi yapılması tercih edilmektedir.



Şekil 3.9 Geçerlilik hatası en düşük düzeyine ulaşıncaya kadar sınıflandırıcı eğitimine devam edilir (Duda et al. 2001).

Bu çalışmada, 3-parçalı çapraz geçerlilik yöntemi ve birisi-hariç yöntemi kullanılarak sınıflandırıcı başarımları değerlendirilmiştir.

BÖLÜM 4

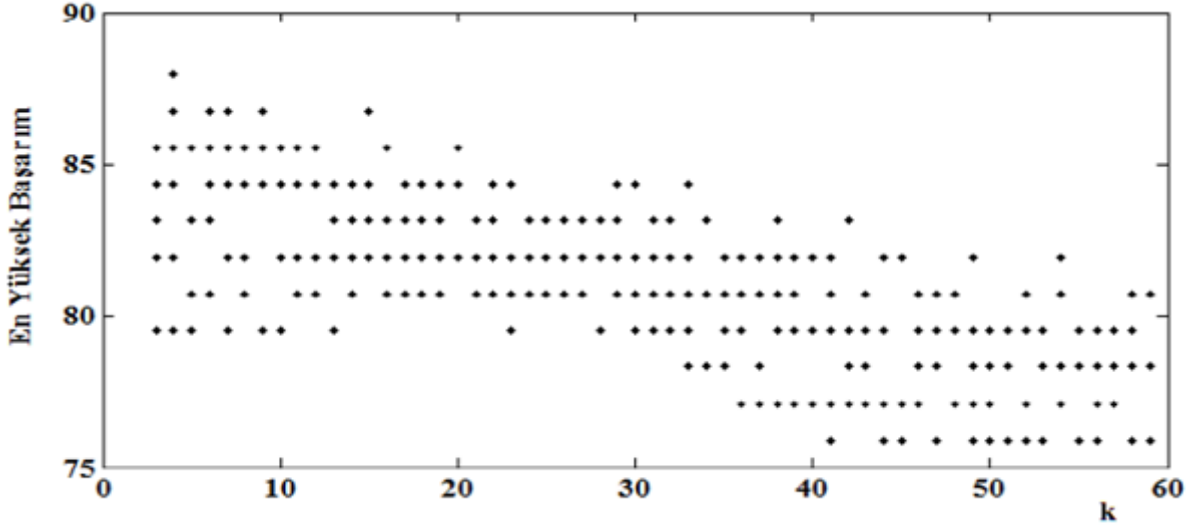
SONUÇLAR

Bu bölümde, elde edilen öznitelikleri kullanarak temel bileşen analizinin teşhise etkisi, k ortalama kümeleme algoritmasının teşhis başarımı, istatistiksel yöntemle öznitelik seçiminin teşhise etkisi, farklı sınıflandırıcıların teşhis başarımı ve son olarak istatistiksel yöntemle öznitelik seçimi sonucunda elde edilen öznitelikleri kullanarak çok katlı sınıflandırıcılarla teşhis başarımı araştırılmış ve sonuçları karşılaştırılmıştır.

4.1 TEMEL BİLEŞEN ANALİZİNİN TEŞHİSE ETKİSİ

Bu çalışma Matlab 2011a yazılımı içerisindeki “kmeans” komutu kullanılarak k-ortalama kümeleyici algoritmasıyla gerçekleştirilmiştir. Ayrıca çalışmada, 59 adet KHD ölçümünden elde edilen veri seti kullanılmıştır. Bu veri seti üzerinde Temel Bileşen Analizi (TBA) farklı anlamlılık düzeyleri (%0,1, %0,5, %1, %5, %10 ve %20) ile uygulanmıştır. Daha sonra, hem farklı düzeylerde TBA'nın kullanıldığı ve hiç kullanılmadığı durumlarda k-ortalama kümeleme algoritmasının hasta ve sağlamı ayırt etme başarımları incelenmiştir. K-ortalama kümeleyicisinin başarımı, seçilen başlangıç konumlarına göre değişebildiği için, her algoritma 100'er defa tekrarlanmıştır. Algoritmada kullanılan küme sayısı (k) 2'den 60'a kadar olan tüm değerler için çalışma gerçekleştirilmiş olup algoritma toplam 82600 defa çalıştırılmıştır. Farklı küme sayılarına karşılık gelen en yüksek algoritma başarımları da kayıt edilerek tek bir grafik halinde sunulmuştur (Şekil 4.1). Biri-dışarıda yöntemiyle test edilerek çalıştırılan k-ortalama kümeleme algoritması sonucunda farklı TBA uygulanma oranlarına göre elde edilen ortalama ve en yüksek başarımları ile TBA sonucu elde edilen öznitelik sayıları Çizelge 4.1'de gösterilmiştir. Tabloya göre en yüksek algoritma başarımı %0,1 oranında değişiminin dışarıda bırakıldığı TBA uygulanması ile elde edilmiştir. Bununla birlikte en yüksek ortalama algoritma başarımı ise %0,5 değişiminin dışarıda bırakılması ile elde edilmiştir. Ayrıca sadece değişiminin %5'ten fazlasının dışarıda bırakılacak şekilde uygulanması halinde en yüksek

algoritma başarımları TBA'nın hiç uygulanmadığı duruma göre daha düşük olmaktadır. TBA uygulandığı her durumda algoritmalarda kullanılan öznitelik sayısı düşmektedir.



Şekil 4.1 K-Ortalama algoritması kullanarak farklı küme sayılarına göre elde edilen en yüksek başarımlar.

Çizelge 4.1 Temel bileşen analizi uygulama oranına göre ortalama sınıflandırıcı başarıları yüzdeleri ve kullanılan öznitelik sayıları

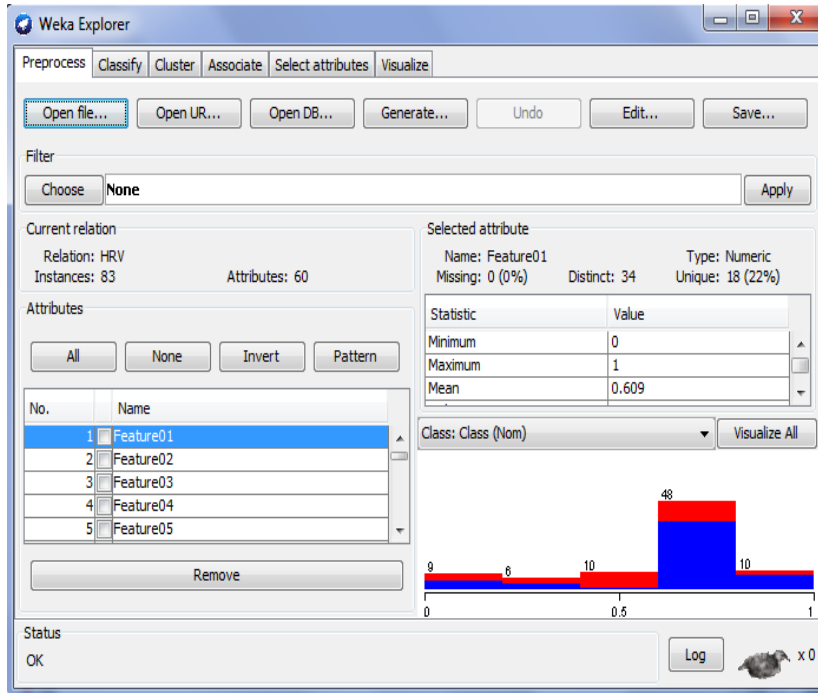
Temel Bileşen Analizi	Ortalama Başarım (%)	En Yüksek Başarım (%)	Öznitelik Sayısı
YOK	74,50	85,54	59
%0,1	74,68	87,95	27
%0,5	75,71	86,74	17
1%	75,00	86,74	12
5%	74,32	85,54	4
10%	72,99	84,34	3
20%	72,74	83,13	2

Bu sonuçlara göre, sadece %0,1 değişiminin dışarıda bırakılmasıyla öznitelik sayısı 59'dan 27'ye düşmekte ve en yüksek başarıma ulaşabilmektedir. %5'den daha az değişiminin dışarıda bırakıldığı şekliyle TBA uygulanarak öznitelikler arası bağımlılıkların giderilmesiyle en yüksek sınıflandırıcı başarımında artış olduğu görülmektedir. Öte yandan yüksek oranda değişiminin dışarıda bırakılması halinde ortalama performansta düşüşe sebep olduğu görülmektedir. Üstelik algoritmada kullanılan küme sayısı parametresinin değeri arttıkça en

yüksek başarımlarda düşüş olduğu da görülmektedir. Temel bileşen analizi ile yapılan öznitelik seçimi işleminde, kullanılacak sınıflandırıcının özellikleri dikkate alınmadığı halde öznitelikler arası ilişkiyi en aza indirdiğinden, sınıflandırıcı başarımını arttırması beklenen bir durumdur (Narin ve İşler 2012).

4.2 K-ORTALAMA KÜMELEME ALGORİTMASININ TEŞHİS BAŞARIMI

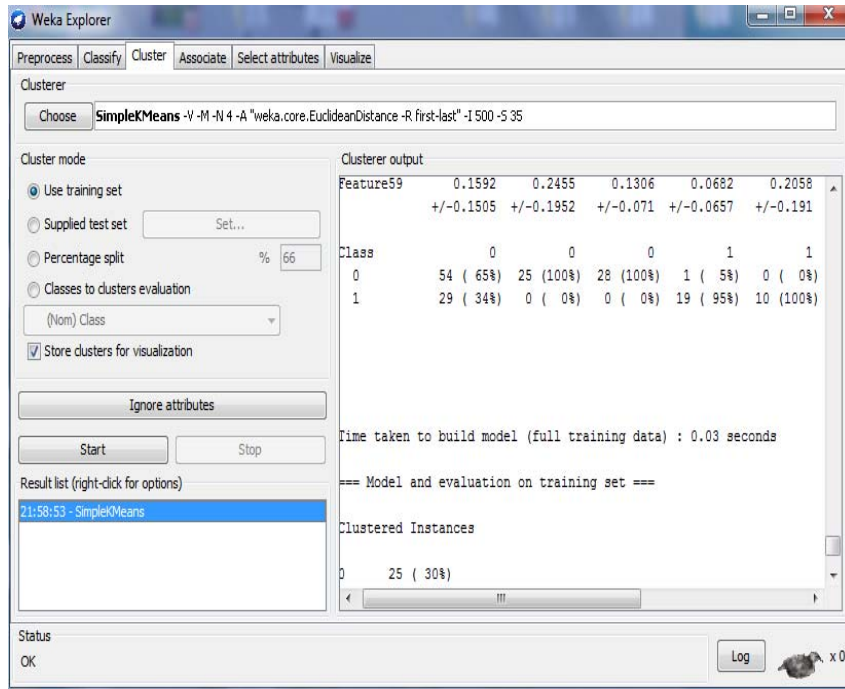
Bu çalışmada, 29'u hasta ve 54'ü hasta olmayan toplam 83 kişiden alınmış 59 adet öznitelikten oluşan KHD veri seti bilgisayar ortamına kayıt edilmiştir. Daha sonra, bu veri seti MATLAB ile geliştirilen bir program kullanılarak WEKA yazılımının desteklediği dosya biçimlerinden biri olan ARFF uzantılı dosya formatında kayıt edilmiştir. WEKA yazılımı kullanılarak "Explorer" ekranından "Open file" komutu ile bu veri dosyası hafızaya alındı. Veri yüklendikten sonra oluşan ekran görüntüsü Şekil 4.2'de gösterilmiştir.



Şekil 4.2 Verinin yüklenmesi sonucu oluşan ara yüz görüntüsü.

Bu ekrandaki "Current relation" başlığı altında yüklenen verinin ismi, kaç tane örneğe ve öznitelige sahip olduğu görülebilmektedir. Verilerden sol tarafta seçili olan öznitelige ait (örnekte Feature01) en küçük, en büyük, standart sapma ve ortalama değer bilgileri ekranda sağ taraftaki "Selected attribute" başlığı altında görülür. Seçilen öznitelige ait dağılım sağ alt köşedeki grafikte görüntülenir. Bu çalışmada kullanılan veri setine herhangi bir ön işlem

uygulanmamıştır. Daha sonra “Cluster” seçeneği seçilip veriye uygulanacak kümeleme algoritması “Clusterer” başlığındaki “Choose” butonu tıklanarak k-Ortalama kümeleme algoritması için “SimpleKMeans” seçilir. Bu seçimden sonra oluşan ekran görüntüsü Şekil 4.3’deki gibidir. Seçilen algoritmanın üstüne tıklanarak algoritmayla ilgili ayarlar yapılmış olup “Choose” butonunun sağındaki metin kutusu içinde bu ayarlar görülmektedir. Bu çalışmada öznitelik seçimi yapılmamış olup eğitim için verilen tüm veriler aynı zamanda test için kullanılmıştır. “Start” aracılığıyla çalıştırılan deney için elde edilen sonuçlar aşağıdaki şekildeki “Clusterer output” başlığı altında görülmektedir.



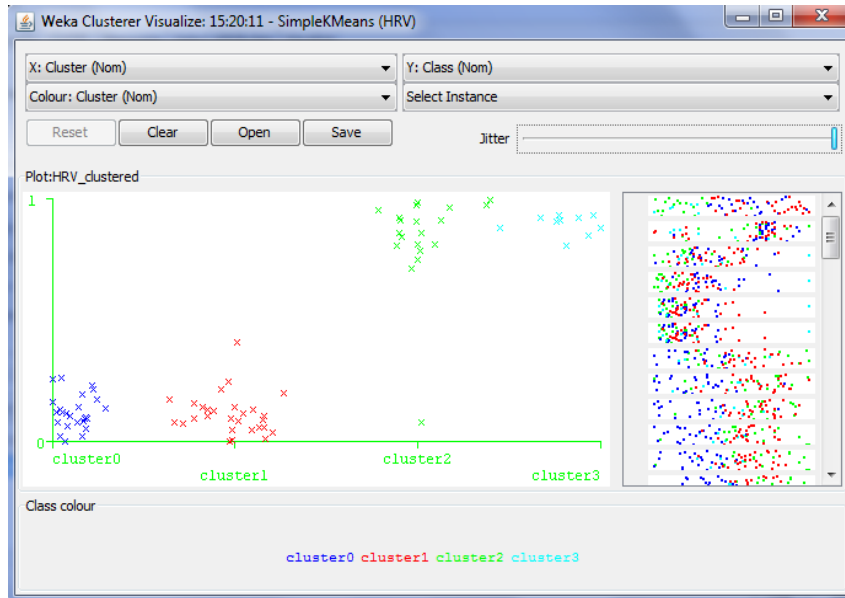
Şekil 4.3 K-ortalama kümeleme algoritması ile kümeleme.

Çalışmada küme sayısını belirleyen k değeri 2’den 10’a kadar alınarak yukarıdaki işlemler farklı başlangıç değerleri (seed) için tekrarlanmıştır. Başlangıç değerinin etkisinin gözlenebilmesi için her k değeri için 1’den 50’ye kadar seed değerleri denenmiştir. Böylece çalışma toplam 450 kez tekrarlanmıştır. Her bir k değeri için elde edilen en yüksek başarımlar ve bu başarımlar için uygulanan seed değerleri ’de özetlenmiştir. Bu sonuçlara göre KKY hastalarının teşhisinde k=4 değeri için maksimum başarımlar %98,79 olarak elde edilmiştir. Böylece 83 kayıttan 82 tanesi doğru olarak tanımlanabilmiştir. WEKA yazılımı “Visualise” sekmesiyle en yüksek başarımların elde edildiği k=4 değeri için sonuçlar görsel olarak elde edilmiştir (Şekil 4.4). Bu şekilde hastalar olmayan birey (0) sınıfı için Cluster0 ve Cluster1 kümeleri, hasta (1) sınıfı için Cluster2 ve Cluster3 kümeleri atanmıştır. Ayrıca sadece 1 adet

hasta olmayan kişinin Cluster2 kümesine dahil edildiği (dolayısıyla hasta olarak belirlendiği) görülmektedir.

Çizelge 4.2 Farklı küme sayısı ve başlangıç değerleri için kümeleyici başarımları

Küme Sayısı (k)	Başarım (%)	Seed (Başlangıç)
2	92,77	3
3	93,97	25
4	98,79	25
5	97,59	5
6	97,59	5
7	96,38	3
8	96,38	5
9	96,38	15
10	96,38	15



Şekil 4.4 En yüksek başarımın elde edildiği 4 adet küme için görselleştirme ekranı.

Çalışmada elde edilen sonuçlara göre KKY hastalığının teşhisinde k-ortalama kümeleyicisinin oldukça başarılı olduğu görülmüştür. Bununla birlikte k-ortalama algoritması kullanılırken

başlangıç merkez değerlerinin algoritma başarımı etkilediği bilindiğinden, farklı başlangıç değerleri için algoritma tekrarlanmış ve en yüksek başarımların verildiği sonuç tablosunda bu değerler de yazılmıştır.

Yapılan bu çalışma, Literatürde çokça tercih edilen Weka 3.6 veri madenciliği programı içindeki k-ortalama kümeleyici algoritması kullanılarak sonuçlar elde edilmiştir (İşler ve Narin 2012).

4.3 FARKLI SINIFLANDIRICILARIN TEŞHİS BAŞARIMININ ANALİZİ

Bu çalışmada kullanılan algoritmalar ve KHD ölçümlerinin hesaplanması için MATLAB 2011a yazılımı kullanılmıştır. Sınıflandırıcı algoritmaları için Matlab yazılımıyla gelen hazır fonksiyonlar kullanılmıştır. Bu çalışmada 29 KKY hastası ve 54 normal birey için oluşan KHD verilerinden hesaplanan 59 adet KHD ölçümü kullanılmıştır. Uygulamada her algoritma 100 defa çalıştırılmıştır. Tüm algoritmaların başarımları 3 parçalı çapraz doğrulama yöntemi kullanılarak elde edilmiştir.

Bu çalışmada kullanılan MLP sınıflandırıcısı 1-50 arasındaki saklı katmandaki nöron sayıları, RBF sınıflandırıcısı 0,1-3,0 arasındaki 0,1 aralıklı yayılım parametre değerleri, SVM sınıflandırıcısı doğrusal çekirdek fonksiyonu, KNN sınıflandırıcısı 1-19 arasındaki tek değerlerdeki k değerleri ve LDA sınıflandırıcısı ise hem doğrusal hem de ikinci dereceden fonksiyonlar için denenmiştir. Elde edilen sınıflandırıcı başarımları Tablo 4.3'de özetlenmiştir. Bu tablodaki en yüksek başarımların değerleri MLP için saklı katman değerinin 11 olduğu durumda, RBF için yayılım parametresinin 1,0 olduğu durumda, KNN için 7 en yakın komşu ve LDA için doğrusal fonksiyonlar kullanıldığında elde edilmiştir.

Çizelge 4.3 5 farklı sınıflandırıcı ile sınıflandırma başarımları. En yüksek başarımlar (ACC), duyarlılık (SEN), özgünlük (SPE), pozitif seçicilik (POS) ve negatif seçicilik (NEG).

Sınıflandırıcılar	Başarımlar				
	ACC	SEN	SPE	POS	NEG
LDA	87,95	82,75	90,74	82,75	90,74
KNN	85,54	65,51	96,26	90,47	83,87
MLP	91,56	82,75	96,29	92,31	91,23
RBF	86,74	79,31	90,74	82,14	89,09
SVM	87,95	72,41	96,29	91,3	86,67

Çizelge 4.3’de, çalışma sonucunda elde edilen en yüksek başarımlar (ACC), duyarlılık (SEN), özgünlük (SPE), pozitif seçicilik (POS) ve negatif seçicilik (NEG) değerleri koyu renkli olarak gösterilmiştir. Genel olarak bakıldığında çok katmanlı algılayıcı (MLP) yapay sinir ağı sınıflandırıcısının en başarılı sınıflandırıcı olduğu görülmektedir.

Bu çalışmada KKY hastalarının normal kontrol grubundan ayırt edilmesinde en yüksek başarımların gizli katmanında 11 nöron bulunan MLP yapay sinir ağı algoritması kullanılarak %91,56 olarak elde edilmiştir. Öznitelik seçimi yapılmamış olmasına rağmen, literatürdeki benzer çalışmalara yakın bir başarımlar elde edilmiştir (Narin vd. 2013a).

4.4 İSTATİSTİKSEL YÖNTEMLE ÖZNETELİK SEÇİMİNİN TEŞHİSE ETKİSİ

Bu çalışmada, öncelikle anormal vurusu olmayan 5 dakikalık KHD verisi bilgisayara yüklenir. İlk olarak, 5 dakikalık bu verileri kullanarak standart KHD ölçümleri, dalgacık dönüşümü ile frekans alanı ölçümleri ve bunlara ek olarak lineer olmayan ölçümlerin eklenmesiyle 59 adet öznitelik elde edilir. İkinci olarak, bu öznitelikleri kullanarak geriye eleme yöntemiyle elenmiş öznitelikleri belirlemek. Son olarak, farklı anlamlılık değerleri için istatistiksel olarak öznitelik seçimi yapılmıştır. Oluşan bu üç farklı öznitelik gruplarını kullanarak, bu grupların sınıflandırıcılardaki başarımlar performansları karşılaştırılmıştır. Sonuçların tekralanabilirliğini sağlamak için her bir sınıflandırıcı farklı parametre değerlerine için 30 defa çalıştırılmıştır. Bu sınıflandırıcılar:

- LDA, doğrusal ve polinomsal kernel fonksiyonları için,
- KNN, k=1, 3, 5, 7, 9, 11, 13 değerleri için,
- MLP, saklı katman nöron sayısının 1–50 arasındaki değerleri için,
- SVM, doğrusal kernel için,
- RBF, 0,1 ile 3,0 arasında 0,1 artırımlı dağılım parametre değerleri için çalıştırılmıştır.

En iyi öznitelikler kombinasyonu ve sınıflandırıcı parametreleri için değerler elde edilmiştir (Çizelge 4.4). Diğer öznitelik seçim yöntemi olan istatistiksel bağımsız örneklem t-testi için 1%, 2%, 5%, 10%, ve 20% anlamlılık değerleri kullanılmıştır. Seçilen öznitelik sayıları sırasıyla 22, 26, 34, 39, ve 48'dir. Bu 5 anlamlılık değeri için seçilen öznitelikler sınıflandırıcılara giriş olarak uygulanmıştır. En iyi öznitelikler kombinasyonu ve sınıflandırıcı parametreleri için değerler elde edilmiştir (Çizelge 4.5).

Çizelge 4.4 Tüm öznitelikler ve geriye doğru eleme yöntemi kullanılarak elde edilen öznitelikleri kullanarak sınıflandırıcı performansının analizi, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznitelik sayısını gösterir.

Sınıflandırıcılar	Tüm özniteliklerle elde edilen başarımlar				Geriye doğru eleme yöntemiyle elde edilen özniteliklerle elde edilen başarımlar			
	SEN	SPE	ACC	NF	SEN	SPE	ACC	NF
1-NN	55,17	77,77	69,87	59	75,86	79,62	78,31	26
3-NN	58,62	87,03	77,10	59	65,51	92,59	83,13	40
5-NN	51,72	92,54	78,31	59	65,51	96,29	85,54	27
7-NN	51,72	96,29	80,72	59	62,06	98,14	85,54	18
9-NN	55,17	96,29	81,92	59	62,06	98,14	85,54	23
11-NN	51,72	96,29	80,72	59	62,06	98,14	85,54	13
13-NN	48,27	98,14	80,72	59	51,72	98,14	81,92	23
Doğrusal LDA	79,31	81,48	80,72	59	79,31	87,03	84,33	33
Polinomsal LDA	68,96	85,18	79,51	59	75,86	90,74	85,54	42
SVM	75,86	85,18	81,92	59	82,75	96,29	91,56	27
MLP	62,06	100,00	86,74	59	82,75	92,59	89,15	18
RBF	65,51	85,18	78,31	59	58,62	96,29	83,13	18

Bu çalışmada, oldukça yaygın kullanılan ve herkesin erişimine açık olan HRV verileri kullanılmış, 59 adet kısa zamanlı KHD ölçümlerinden elde edilmiş özniteliklerin farklı kombinasyonları arasındaki sınıflar arası ayırım başarımları hem geriye doğru eleme hem 1%,

2%, 5%, 10%, ve 20% anlamlılık deęerleri iin istatistiksel baęımsız rneklem t-testi kullanılarak incelenmiřtir.

Seilen her znelik kumesinin performansını deęerlendirmek iin literatürde ok yaygın kullanılan 5 adet sınıflandırıcı algoritması kullanılmıřtır.

Çizelge 4.5 %1, %2, %5, %10 ve %20 istatistiksel anlamlılık değerleri için öznitelik seçimi ile sınıflandırıcı başarımları, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznitelik sayıdır.

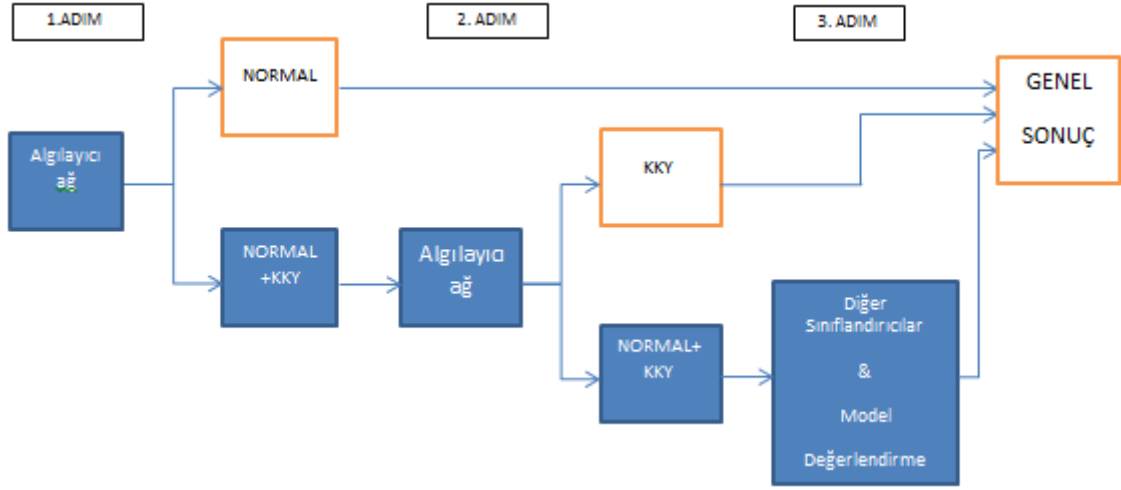
Sınıflandırıcılar	p=%1 (NF=22)			p=%2 (NF=26)			p=%5 (NF=34)			p=%10 (NF=39)			p=%20 (NF=48)		
	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC
1-NN	62,06	85,18	77,10	68,96	83,33	78,31	55,17	83,33	73,49	58,62	87,03	77,10	48,27	79,62	68,67
3-NN	65,51	92,59	83,13	65,51	94,44	84,33	62,06	94,44	83,33	58,62	88,88	78,31	65,51	85,18	78,31
5-NN	55,17	94,44	80,72	58,62	94,44	81,92	58,62	94,44	81,92	58,62	90,74	79,51	51,72	94,44	79,51
7-NN	55,17	92,59	79,51	58,62	94,44	81,92	58,62	96,26	83,13	48,27	96,29	79,51	48,27	96,29	79,51
9-NN	55,17	92,59	79,51	55,17	92,59	79,51	58,62	96,29	83,13	55,17	96,29	81,92	51,72	96,29	80,72
11-NN	55,17	90,74	78,31	55,17	92,59	79,51	55,17	96,29	81,92	55,17	96,29	81,92	51,72	96,29	80,72
13-NN	48,27	92,59	77,10	55,17	92,59	79,51	55,17	96,29	81,92	51,72	96,29	80,72	51,72	96,14	81,92
Doğrusal LDA	75,86	81,48	79,51	78,86	83,33	80,72	79,31	85,18	83,13	79,31	83,33	81,92	79,31	77,77	78,31
Polinomsal LDA	68,96	85,18	79,51	72,41	87,03	81,92	72,41	88,88	83,13	72,41	87,03	81,92	68,96	87,03	80,72
MLP	75,86	92,59	86,74	75,86	92,59	86,74	79,31	92,59	87,95	68,96	96,29	86,74	68,96	94,44	85,54
SVM	72,41	81,48	78,31	62,06	85,18	77,10	72,41	90,74	84,33	65,51	87,03	79,51	62,06	83,33	75,90
RBF	48,27	94,44	78,31	75,86	75,92	75,90	41,37	94,44	75,90	58,62	90,74	79,51	65,51	87,03	79,51

Çizelge 4.4 ve 4.5’de elde edilen sonuçlara göre öznelik seçimi için kullanılan her iki yöntemde de algoritmalarımızın performans başarımlarını arttırmıştır. En yüksek performans başarımları geriye doğru eleme yöntemi kullanılarak elde edilmiştir. Bu değerler hassaslık (SEN), özgünlük (SPE) ve doğruluk (ACC) olarak sırasıyla 82,75%, 96,29%, ve 91,56% olarak bulunmuştur. Bu değerler literatürde oldukça yüksek değerlerdir.

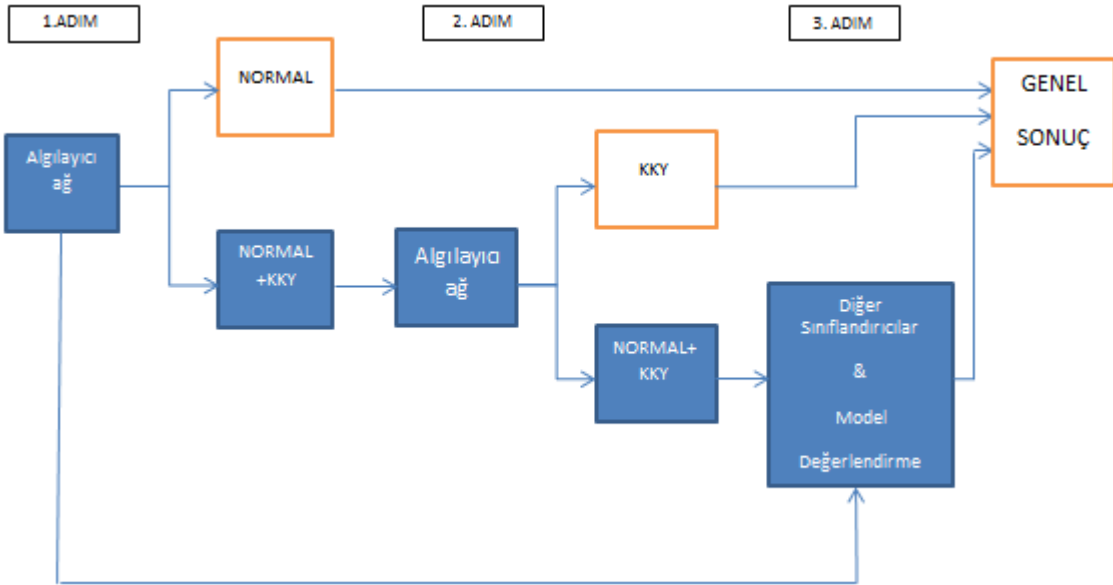
Maksimum performans değeri 27 öznelik seçilmiş SVM kullanılarak bulunmuştur. Bu çalışmadaki en önemli nokta lineer olmayan ölçümler ve dalgacık dönüşümü ile elde edilen frekans alanı ölçümlerini içeren özneliklerin oldukça iyi başarımlar vermesidir. Bu ölçümlerin klasik KHD öznelikleri içinde anlamlı düzeyde KKY hastalığının teşhisinde kullanılabilir olduğunu elde edilen sonuçlardan görülmektedir (Narin vd. 2013b).

4.5 ÇOK KATLI SINIFLANDIRICILARLA TEŞHİS BAŞARIMI

Bu çalışmada 29 KKY rahatsızlığı bulunan ve 54 normal bireylerden alınmış kısa süreli KHD verilerinden elde edilen 59 adet özneliği kullanarak çok katmanlı sınıflandırıcı sistemi tasarlanmıştır (Şekil 4.5, Şekil 4.6, Şekil 4.7, Şekil 4.8). Burada, Şekil 4.5 ve 4.6 için ilk olarak elde edilen veriler algılayıcı ağa giriş olarak verilir. Buradan normal yani hasta olmayan bireylerin en yüksek sayıda doğru olarak tespit edilenleri ayrılır. Varsa kalan normal bireyler ve KKY rahatsızlığı olan bireyler tekrardan algılayıcı ağa giriş olarak gönderilir. Burada ise KKY bireylerin en fazla doğru olarak tespit edilenleri ayrılır. Varsa kalan KKY ve bir önceki adımda kalan normal bireyler MLP, SVM, RBF, KNN, LDA algoritmalarına ayrı ayrı giriş verisi olarak uygulanır. Sonuç olarak son aşamadaki algoritmaların başarımları ve önceki algılayıcı ağ çıkışlarındaki başarımlar yorumlanarak sistemin genel sonucu bulunur. Burada başarımların değerlendirilmesi iki yöntemle bulunur. Birinci yöntemde, 3. Aşamaya gelen bireylerin kendi içinde birini dışarda bırak çapraz doğrulama yöntemini kullanarak sınıflandırma başarımının bulunması ve bulunan bu değerlere 1. ve 2. aşamalarda elde edilen sonuçlar eklenerek genel sonuç bulunmaktadır (Şekil 4.5). İkinci yöntemde, 3. aşamaya gelen bireylerin tümünün, tüm veri kümesi üzerinde test verisi olarak kullanılmasıyla elde edilen sınıflandırma başarımına 1. ve 2. aşamalarda elde edilen sonuçlar eklenerek genel sonuç bulunmaktadır (Şekil 4.6).



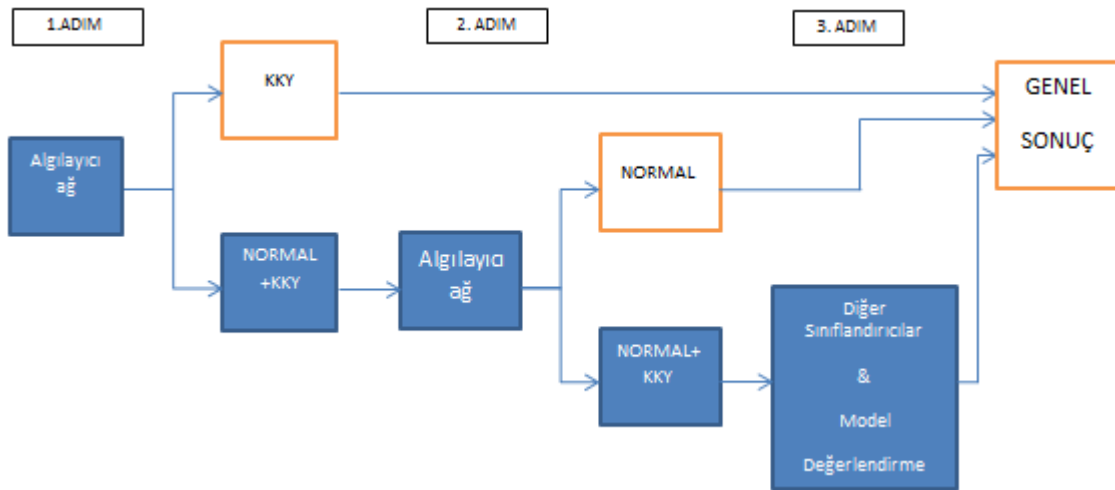
Şekil 4.5 KKY minimum olduğu başlangıç değeri için diğer sınıflandırıcılardan geriye kalan kalan kişilerin birini hariç tut çapraz doğrulama yöntemiyle elde edilen sınıflandırıcı sistemi.



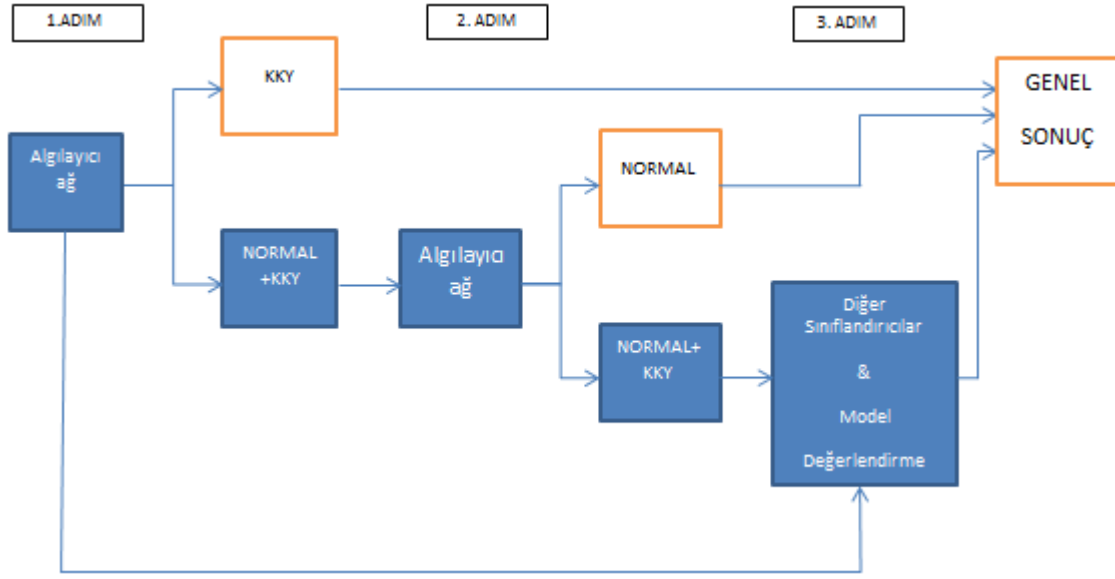
Şekil 4.6 KKY minimum olduğu başlangıç değeri ile tüm veri üzerinden diğer sınıflandırıcılardan geriye kalan kalan kişilerin test amaçlı kullanılması ile elde edilen sınıflandırıcı sistemi.

Bir diğer yaklaşım, Şekil 4.7 ve 4.8 için eldeki veriler algılayıcı ağı giriş olarak verilir. Buradan KKY rahatsızlığı olan bireylerin en yüksek sayıda doğru olarak tespit edilenleri ayrılır. Varsa kalan KKY rahatsızlığı olan bireyler ve normal bireyler tekrardan algılayıcı ağı giriş olarak gönderilir. Burada ise hasta olmayan yani normal bireylerin en fazla doğru olarak tespit edilenleri ayrılır. Varsa normal bireyler ve bir önceki adımda kalan KKY rahatsızlığı

olan bireyler MLP, SVM, RBF, KNN, LDA algoritmalarına ayrı ayrı giriş verisi olarak uygulanır. Sonuç olarak son aşamadaki algoritmaların başarımları ve önceki algılayıcı ağ çıkışlarındaki başarımlar yorumlanarak sistemin genel sonucu bulunur. Burada başarımların değerlendirilmesi iki yöntemle bulunur. Birinci yöntemde, 3. Aşamaya gelen bireylerin kendi içinde birini dışarda bırak çapraz doğrulama yöntemini kullanarak sınıflandırma başarımının bulunması ve bulunan bu değerlere 1. ve 2. aşamalarda elde edilen sonuçlar eklenerek genel sonuç bulunmaktadır (Şekil 4.7). İkinci yöntemde, 3. aşamaya gelen bireylerin tümünün, tüm veri kümesi üzerinde test verisi olarak kullanılmasıyla elde edilen sınıflandırma başarımına 1. ve 2. aşamalarda elde edilen sonuçlar eklenerek genel sonuç bulunmaktadır (Şekil 4.8). Bu iki sistem doğrultusunda elde edilen bu sonuçlar ayrıntılı olarak Çizelge 4.6 ve Çizelge 4.7 de bulunmaktadır. Çizelge 4.6 ve 4.7'deki %1, %2, %5, %10 ve %20 değerleri istatistik anlamlılık değerlerini, KKY_min ifadesi 1. Adımda algılayıcı ağın başlangıcında KKY rahatsızlığı olan hastaların minimum hata ile sınıflandırılmasını yani en yüksek düzeyde doğru olan KKY rahatsızlığına sahip bireyleri, Normal_min ifadesi ise 1. adımda algılayıcı ağın başlangıcında hasta olmayan yani normal bireylerin minimum hata ile sınıflandırılmasını yani en yüksek düzeyde doğru olan normal bireylerin başlangıç kriteri olarak belirlenmesini temsil etmektedir. NF değeri ise istatistiksel anlamlılık değerleri sonucunda elde edilen öznelik sayısını gösterir. Çizelgelerdeki “yok” ifadesi ise 3. aşamaya yeterli sayıda bireyin kalmaması nedeniyle 3. Aşamada sonuç üretmemesini temsil etmektedir.



Şekil 4.7 Normal bireylerin minimum olduğu başlangıç değeri için diğer sınıflandırıcılardan geriye kalan kişilerin birini hariç tut çapraz doğrulama yöntemiyle elde edilen sınıflandırıcı sistemi.



Şekil 4.8 Normal bireylerin minimum olduğu başlangıç değeri için tüm veri üzerinden algılayıcı sınıflandırıcısından geriye kalan kalan kişilerin test amaçlı kullanılması ile elde edilen sınıflandırıcı sistemi.

Çizelge 4.6 ve 4.7’den elde edilen sonuçlara göre 1%, 2%, 5%, 10%, ve 20% anlamlılık değerleri için bağımsız örneklem t-testi kullanılarak öznitelik seçimi yapılmış veriler kullanılmıştır. En yüksek performans başarımı Çizelge 4.7’de görüldüğü gibi, son adımda elde edilen bireylerin kendi içinde birini hariç tut çapraz doğrulama yöntemiyle gerçekleştirilmiş işlemin sonucunda elde edilmiştir. Bu değerler hassaslık (SEN), özgünlük (SPE) ve doğruluk (ACC) olarak sırasıyla 100.00%, 100.00%, ve 100.00% olarak bulunmuştur. Bu değerler literatürdeki en yüksek değerlerdir.

Bu maksimum performans değeri 5% anlamlılık değerinde 34 öznitelik seçilerek ve algılayıcı ağın başlangıç kriteri normal bireylerin minimum hata ile elde edilmesiyle bulunmuştur. Elde edilen öznitelikler ve başlangıç kriterlerinde MLP sınıflandırıcı algoritmasında en yüksek değerler elde edilmiştir. Elde edilen bu değerler sonucunda çok katmanlı bir sınıflandırıcı sistemi tasarlanarak sınıflandırıcı başarımlarının artacağı tespit edilmiş. Böyle sistemlerin hem zaman hem de son adımda kullanılan hesapsal olarak yoğun işlem gerektiren algoritmalarda işlem yükünü azaltacağı görülmektedir. Günümüzde buna benzer bir uygulamayı şu anki pratisyen hekim ve uzman hekim olarak değerlendirebiliriz. Yani pratisyen hekim bizim algılayıcı ağımızı, uzman hekim ise kuvvetli sınıflandırıcı algoritmalarını temsil etmektedir. Bu nedenle, böyle bir yapıda karmaşa ve zaman israfından çok fazla söz edilemez. Bununla beraber daha yüksek başarımlar ve daha güvenilir sonuçlar elde edilmektedir.

Çizelge 4.6 %1, %2, %5, %10 ve %20 istatistik anlamlılık değerleri için öznitelik seçimi ile elde edilen öznitelikler kullanılarak tüm veri üzerinden algılayıcı sınıflandırıcısından geriye kalan kalan kişilerin test amaçlı kullanılması ile elde edilen sınıflandırıcı başarımları, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznitelik sayını, KKY_min konjestif kalp yetmezliği hastalarının minumum hata ile sınıflandırılması, NORMAL_min normal bireylerin minumum hata ile sınıflandırılmış olduğu başlangıç değerlerini gösterir.

Algoritmalar	Perceptron başlangıç kriteri	1%			2%			5%			10%			20%			100%		
		SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC
LDA	KKY_min	75,86	96,29	89,15	72,41	100,00	90,36	82,75	100,00	93,97	82,75	100,00	93,97	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	82,75	79,62	80,72	93,10	86,18	87,95	93,10	96,29	95,18	96,55	88,88	91,56	100,00	87,03	91,56	100,00	85,18	90,36
MLP	KKY_min	82,75	98,14	92,77	82,75	100,00	93,97	100,00	98,14	98,79	100,00	98,14	98,79	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	75,86	98,14	90,36	86,20	100,00	95,18	93,10	100,00	97,59	93,10	100,00	97,59	96,55	100,00	98,79	96,55	100,00	98,79
KNN	KKY_min	75,86	96,29	89,15	75,86	100,00	91,56	79,31	100,00	92,77	79,31	100,00	92,77	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	86,20	72,22	77,10	93,10	74,07	80,72	89,65	98,14	95,18	93,10	94,44	93,97	96,55	92,59	93,77	96,55	94,44	95,18
SVM	KKY_min	68,96	98,14	87,95	72,41	100,00	90,36	79,31	100,00	92,77	79,31	100,00	92,77	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	82,75	70,37	74,69	93,10	51,85	66,26	93,10	88,88	90,36	93,10	90,74	91,56	100,00	85,18	90,36	100,00	85,18	90,36
RBF	KKY_min	75,86	96,29	89,15	72,41	100,00	90,36	100,00	98,14	98,79	100,00	98,14	98,79	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	72,41	94,44	86,74	96,55	64,81	75,90	89,65	100,00	96,38	93,10	100,00	97,59	96,55	100,00	98,79	96,55	100,00	98,79
Öznitelikler(NF)		22			26			34			39			48			59		

Çizelge 4.7 %1, %2, %5, %10 ve %20 istatistik anlamlılık değerleri için öznitelik seçimi ile elde edilen öznitelikler kullanılarak algılayıcı sınıflandırıcısından geriye kalan kişilerden birini hariç tut çapraz doğrulama yöntemiyle elde edilen sınıflandırıcı başarımları, SEN hassalık değerini (%), SPE özgünlük değerini (%), ACC genel doğruluk değerini (%) ve NF ise öznitelik sayını, KKY_min konjestif kalp yetmezliği hastalarının minimum hata ile sınıflandırılması, NORMAL_min normal bireylerin minimum hata ile sınıflandırılmış olduğu başlangıç değerlerini gösterir.

Algoritmalar	Perceptron başlangıç kriteri	1%			2%			5%			10%			20%			100%		
		SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC
LDA	KKY_min	82,75	92,59	89,15	86,20	96,29	92,77	93,10	98,14	96,38	93,10	98,14	96,38	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	79,31	88,88	85,54	86,20	94,44	91,56	89,65	100,00	96,38	96,55	96,29	96,38	96,55	98,14	97,59	96,55	94,44	95,18
MLP	KKY_min	89,65	96,29	93,97	89,65	98,14	95,18	100,00	98,14	98,79	100,00	98,14	98,79	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	89,65	98,14	95,18	89,65	100,00	96,38	100,00	100,00	100,00	96,55	100,00	98,79	96,55	100,00	98,79	96,55	100,00	98,79
KNN	KKY_min	68,96	100,00	89,15	79,31	98,14	91,56	100,00	98,14	98,79	100,00	98,14	98,79	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	72,41	100,00	90,36	86,20	100,00	95,18	89,65	100,00	96,38	93,10	100,00	97,59	96,55	100,00	98,79	96,55	100,00	98,79
SVM	KKY_min	79,31	88,88	85,54	86,20	94,44	91,56	yok	yok	yok	yok	yok	yok	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	82,75	88,88	86,74	86,20	94,44	91,56	89,65	96,29	93,97	96,55	96,29	96,38	Yok	Yok	Yok	Yok	Yok	Yok
RBF	KKY_min	89,65	92,59	91,56	82,75	96,29	91,56	100,00	98,14	98,79	100,00	98,14	98,79	Yok	Yok	Yok	Yok	Yok	Yok
	NORMAL_min	82,75	96,29	91,56	86,20	100,00	95,18	89,65	100,00	96,38	93,10	100,00	97,59	96,55	100,00	98,79	96,55	100,00	98,79
Öznitelikler		22			26			34			39			48			59		

BÖLÜM 5

TARTIŞMA

Elde edilen tüm bu sonuçlara göre, ilk olarak sadece %0,1 değişiminin dışarıda bırakılmasıyla değişken sayısı 59'dan 27'ye düşmekte ve en yüksek başarıma ulaşılabilmektedir. %5'den daha az değişiminin dışarıda bırakıldığı şekliye TBA uygulanarak öznitelikler arası bağımlılıkların giderilmesiyle en yüksek sınıflandırıcı başarımında artış olduğu görülmektedir. Öte yandan yüksek oranda değişiminin dışarıda bırakılması halinde performans artışı yerine düşüşe sebep olduğu görülmektedir. Üstelik algoritmada kullanılan küme sayısı parametresinin değeri arttıkça en yüksek başarımlarda düşüş olduğu da görülmektedir. Temel bileşen analizi ile yapılan öznitelik seçimi işleminde, kullanılacak sınıflandırıcının özellikleri dikkate alınmadığı halde öznitelikler arası ilişkiyi en aza indirdiğinden, sınıflandırıcı başarımını arttırması beklenen bir durumdur. Bununla birlikte, KHD hastalarının teşhisi çalışmalarında temel bileşen analizinin yanı sıra diğer genetik algoritma gibi yapay zeka tabanlı öznitelik seçimi algoritmalarının kullanılması halinde başarımın daha fazla artması beklenmektedir.

İkinci olarak, günümüzde inanılmaz oranda artan veri miktarından dolayı istenen bilgiye kolayca ulaşabilme ve veriden anlamlı sonuçlar elde edebilme zorlaşmıştır. Bunun daha kolay gerçekleştirilmesi için yazılımlardan faydalanılmaktadır. WEKA yazılımı ile yapılmış bu çalışmada konjestif kalp yetmezliği hastalarının teşhisi uygulamasında kullanılmıştır. Diğer yazılımlara göre kolay elde edilebilir ve açık kaynak kodlu olması sebebiyle kullanıcı sayısı fazladır. WEKA, içerisinde bulundurduğu birçok algoritma sayesinde verilerden bilgi çıkarımı yapabilmesine olanak sağlamaktadır.

Çalışmada elde edilen sonuçlara göre KKY hastalığının teşhisinde k ortalama kümeleyicisinin oldukça başarılı olduğu görülmüştür. Bununla birlikte k-ortalama algoritması kullanılırken başlangıç merkez değerlerinin algoritma başarımı etkilediği bilindiğinden, farklı başlangıç değerleri için algoritma tekrarlanmış ve en yüksek başarımların verildiği sonuç tablosunda bu

değerler de yazılmıştır. WEKA yazılımı kullanılarak gerçekleştirilecek diğer çalışmalarda da farklı başlangıç değerlerinin denenmesi gereklidir.

Çalışmada öznitelik seçimi yapılmamıştır. Öznitelik seçimi yapılan diğer algoritmaların da eklenmesi ve test kümesinin eğitim kümesi dışından seçimine olanak sağlayan diğer algoritmaların da uygulanmasıyla daha güvenilir sonuçların elde edilebileceği ön görülmektedir.

Üçüncü olarak, çalışmada 3 parçalı çapraz doğrulama yöntemiyle gerçekleştirilen model değerlendirmesi sonucunda öznitelik seçimi yapılmamış olmasına karşın oldukça yüksek sonuçlar elde edilmiştir. Sonuçların daha kararlı ve güvenilir olması için ise birisi-dışarıda (leave-one-out) yaklaşımıyla sonuçların değerlendirilmesi çalışma sonucunun sağlığını artırması aynı zamanda öznitelik seçiminde eklendiğinde başarımın, hem daha güvenilir hem daha yüksek olması öngörülmektedir.

Dördüncü olarak, kısa süreli KHD ölçümlerinden elde edilen 59 adet öznitelik değerlerini kullanarak yapılan geriye doğru eleme yöntemiyle öznitelik seçimi ve istatistiksel olarak 1%, 2%, 5%, 10%, ve 20% anlamlılık değerlerini kullanarak yapılan öznitelik seçimi sonucunda elde edilen özniteliklerle çokça kullanılan 5 adet sınıflandırıcı algoritmasının performans analizi yapılmıştır. Sonuçta elde edilen değerler çerçevesinde başarım performansı oldukça artmıştır. Özellikle iki öznitelik seçimini karşılaştırdığımızda geriye doğru eleme yöntemi daha iyi sonuç vermiştir. Özellikle lineer olmayan öznitelik ölçümleri ve dalgacık dönüşümü frekans ölçümleri oldukça iyi sonuçlar vermiştir. Bu gösteriyor ki bu ölçümlerin de dâhil olduğu öznitelikler oldukça yüksek sonuçlar vermektedir. Elde edilen sonuçların başarımlarının artması için uzun süreli KHD verilerinden elde edilen öznitelikler kullanılabilir. Çünkü kısa süreli KHD verileri kalbin günlük aktivitelerin tamamını göstermemektedir.

Son olarak, elde edilen 59 adet öznitelikten ve 83 kişiden oluşan veri kümesini kullanarak öznitelik seçimine gidilmiştir. Öznitelik seçiminde 1%, 2%, 5%, 10%, ve 20% anlamlılık değerlerine sahip bağımsız örneklem t-testi kullanılmıştır. Sonuç olarak elde edilen yeni veri kümesini giriş olarak kullanan çok katlı sınıflandırıcı tasarımı yapılmış. Yapılan bu sistemin başarımı oldukça arttırdığı elde edilen sonuçlardan gözükmektedir. %100 başarımlar sağlanmıştır.

Bu çerçeve içerisinde, ileriki çalışmalarda sonuçların daha hızlı ve daha güvenilir olması için genetik algoritma tabanlı algoritmalarla öz nitelik seçiminin yapılması hedeflenmektedir.

KAYNAKLAR

- Akselrod S, Gordon D, Ubel F A, Shannon D C, Barger A C and Cohen R J** (1981) Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, 213, 220–222.
- Albert N M** (2000) Implementation strategies to manage heart failure outcomes: Managing clinical and economic outcomes. *AACN Clin Issues*, 11(3): 396–411 and 475–477.
- American Heart Association** (2006) Heart Disease and Stroke Statistics–2006 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 113: 85–151.
- Antila K J, Valimaki I A T, Makela M, Tuominen J, Wilson A J and Southall D P** (1990) Heart rate variability in infants subsequently suffering sudden infant death syndrome (SIDS). *Early Hum Dev*, 22: 57–72.
- Artinian N T, Harden J K, Kronenberg M W, Vander Wal J S, Daher E, Stephens Q and Bazzia R I** (2003) Pilot study of a web-based compliance monitoring device for patients with congestive heart failure. *Heart Lung*, 32(4):226–233.
- Asyali M H** (2003) Discrimination power of long-term heart rate variability measures. *In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Cancun.
- Bakardjian H and Yamamoto K** (1995) Dynamic non-deterministic characterization of HRV through multiresolution wavelet decomposition. *In Annual International Conference of the IEEE Engineering in Medicine and Biology – Proceedings*, 17(2): 1063–1064.
- Barbieri R, Triedman J K and Saul J P** (2002) Heart rate control and mechanical cardiopulmonary coupling to assess central volume: a systems analysis. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 283: R1210–R1220.
- Barrett C J, Navakatikyan M A and Malpas S C** (2001) Long-term control of renal blood flow: what is the role of the renal nerves? *Am J Physiol-Reg I*, 280: R1534–R1545.
- Berne R and Levy M** (1997) *Cardiovascular Physiology*. 7th ed.
- Berntson G G, Bigger T, Eckberg D L, Grossman P, Kaufmann P G, Malik M, Nagaraja H N, Porges S W, Saul J P, Stone P H and van der Molen M W** (1997) Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, 34: 623–648.

KAYNAKLAR (devam ediyor)

- Bigger J T, Albrecht P, Steinman R C, Rolnitzky L M, Fleiss J L and Cohen R J** (1989) Comparison of time and frequency domain–based measures of cardiac parasympathetic activity in Holter recordings after myocardial infarction. *Am J Cardiol*, 64: 536–538.
- Bigger J T, Fleiss J L, Steinman R C Rolnitzky L M, Kleiger R E and Rottman J N** (1992) Correlations among time and frequency domain measures of heart period variability two weeks after acute myocardial infarction. *Am J Cardiol*, 69: 891–898.
- Braga A N, da Silva Lemos M, da Silva J R, Fontes W R and dos Santos R A** (2002) Effects of angiotensins on day–night fluctuations and stress–induced changes in blood pressure. *Am J Physiol-Reg I*, 282: R1663–R1671.
- Braune H J and Geisenorfer U** (1995) Measurement of heart rate variations: influencing factors, normal values and diagnostic impact on diabetic autonomic neuropathy. *Diabetes Res Clin Pr*, 29: 179–187.
- Brennan M, Palaniswami M and Kamen P** (2001) Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability? *IEEE T Bio-Med Eng*, 48(11): 1342–1347.
- Byun H and Lee S-W** (2003) A survey of pattern recognition applications of support vector machines. *Int J Pattern Recogn*, 17(3): 459–486.
- Cao H** (2004) *Linear and nonlinear approaches to heart rate analysis near the time of birth*. Ph.D. thesis, University of Virginia, Department of Biomedical Engineering, USA.
- Chang K L Monahan K J, Griffin M P, Lake D E and Moorman J R** (2001) Comparison and clinical application of frequency domain methods in analysis of neonatal heart rate time series. *Ann Biomed Eng*, 29: 764–774.
- Chiarugi F, Colantonio S, Emmanouilidou D, Moroni D and Salvetti O** (2008) Biomedical signal and image processing for decision support in heart failure. *Comput Cardiol*, 35: 649–652.
- Clark D O, Tu W, Weiner M and Murray M D** (2003) Correlates of health–related quality of life among lower–income, urban adults with congestive heart failure. *Heart Lung*, 32(6): 391–401.
- Clifford G D, Azuaje F and McSharry P E** (2006) *Advanced Methods and Tools for ECG Data Analysis*. Artech House, Inc., Norwood, MA.
- Clifford G D and Tarassenko L** (2005) Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. *IEEE T Bio-Med Eng*, 52(4): 630–638.
- Conratha C E and Opthof T** (2005) The Patient U Wave. *Cardiovasc Res*, 67(2): 184- 186.
- Contreras P, Canetti R and Migliaro E R** (2007) Correlations between frequency-domain HRV indices and lagged Poincare plot width in healthy and diabetic subjects. *Physiol Meas*, 28: 85–94.

KAYNAKLAR (devam ediyor)

- Cooley J W and Tukey J W** (1965) An algorithm for the machine computation of complex Fourier series. *Mathematical Computations*, 19: 297–301.
- Costa M, Goldberger A L and Peng C K** (2002) Multiscale entropy analysis of physiologic time series. *Phys Rev Lett*, 89: 062102.
- Costa M, Goldberger A L and Peng C K** (2005) Multiscale entropy analysis of biological signals. *Phys Rev E*, 71: 021906.
- Costa O, Lago P, Rocha A P, Carvalho M J, Freitas A, Freitas J, Puig J, Brando A and de Freitas F** (1994) Heart rate variability in 24-hour Holter recordings. Comparative study between short- and long-term time and frequency-domain analyses. *J Electrocardiol*, 27: 251–254.
- Di Bari M, Pozzi C, Cavallini M C, Innocenti F, Baldereschi G, De Alfieri W Antonini E Pini R, Masotti G and Marchionni N** (2004) The Diagnosis of Heart Failure in the Community Comparative Validation of Four Sets of Criteria in Unselected Older Adults: The ICARe Dicomano Study. *J Am Coll Cardiol*, 44(8): 1601-1608.
- Dionne I, White M and Tremblay M** (2002) The reproducibility of power spectrum analysis of heart rate variability before and after a standardized meal. *Physiol Behav*, 75(3): 267–270.
- Duda R O, Hart P E and Stork D G** (2001) *Pattern Classification*. New York: John Wiley and Sons, 2. baskı.
- Eberhart R C and Dobbins R W** (1990) *Neural network PC tools*. San Diego: Academic Press.
- Edlinger G, Litscher G and Pfurtscheller G** (1994) Analysis of cardio-respiratory signals – methodology and applications in infants. *Biomed Tech*, 39: 274–278.
- Efe E, Bek Y ve Şahin M** (2000) *İstatistik Yöntemler II*, Kahramanmaraş Sütçü İmam Üniversitesi, Bilgisayar Araştırma ve Uygulama Merkezi, Yayın No: 10.
- Ewing D J, Martin C N, Young R J and Clarke B F** (1985) The value of cardiovascular autonomic function tests: 10 years' experience in diabetes. *Diabetes Care*, 8: 491–498.
- Flavell C and Stevenson L W** (2001) Take Heart with Heart Failure. *Circulation*, 104: 89.
- Fukunaga K** (1990) *Introduction to Statistical Pattern Recognition*, 2nd Edition, Academic Press.
- Gibbons R J, Balady G J, Beasley J W, Bricker J T, Duvernoy W F, Froelicher VF, Mark D B Marwick T H, McCallister B D, Thompson P D J, Winters W L, Yanowitz F G, Ritchie J L, Cheitlin M D, Eagle K A, Gardner T J, Garson A J, Lewis R P, O'Rourke R A and Ryan TJ** (1997) ACC/AHA guidelines for exercise testing, a report of the American College of Cardiology/ American Heart Association Task Force on practice guidelines (committee on exercise testing *J Am Coll Cardiol*, 30(1): 260–311.

KAYNAKLAR (devam ediyor)

- Goldberg D E** (1989) *Genetic algorithms in search, optimization, and machine learning*. New York: Addison Wesley Pub. Comp. Inc.
- 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 and Stanley H E** (2000) Physiobank, physiotookit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23): 215-220.
- Goldstein B, Fiser D H, Kelly M M, Mickelsen D, Ruttimann U and Pollack M M** (1998) Decomplexification in critical illness and injury: relationship between heart rate variability, severity of illness, and outcome. *Crit Care Med*, 26: 352–357.
- Grossman P** (1992) Breathing rhythms of the heart in a world of no steady state: a comment on weber, molenaar, and van der molen. *Psychophysiology*, 29(1): 66–72.
- Gura M T and Foreman L** (2004) Cardiac resynchronization therapy for heart failure management. *AACN Clin Issues*, 15(3): 326–339.
- Guyon I and Elisseeff A** (2003) An Introduction to Variable and Feature Selection. *J Mac Learn Res*, 3: 1157- 1182.
- Hamilton R, Mckenchnie P and Macfarlane P** (2004) Can cardiac vagal tone be estimated from the 10-second ECG? *Int J Cardiol*, 95(1): 109–115.
- Harper R M, Hoppenbrouwers T, Sterman M B, McGinty D J and Hodgman J** (1976) Polygraphic studies of normal infants during the first six months of life. I: Heart rate and variability as a function of state. *Pediatr Res*, 10: 945–951.
- Hirsh J A and Bishop B** (1981) Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. *Am J Physiol*, 241: 620–629.
- Hobbs F D R, Korewicki J, Cleland J G F, Eastaugh J and Freemantle N** (2005) The diagnosis of heart failure in European primary care: The IMPROVEMENT programme survey of perception and practice. *Eur J Heart Fail*, 7: 768–779.
- Holland J** (1975) *Adaptation in Natural and Artificial Systems*. Univ. of Michigan Press, Ann Arbor, Mich.
- Hon E H and Lee S T** (1965) Electronic evaluations of the fetal heart rate patterns preceding fetal death: further observations. *Am J Obstet Gynecol*, 814–826.
- Hu K, Ivanov P C, Chen Z, Carpena P and Stanley H E** (2001) Effect of trends on detrended fluctuation analysis. *Phys Rev E*, 64: 1–19.
- Huikuri H V, Makikallio T H, Raatikainen P, Perkiomaki J, Castellanos A and Myerburg R J** (2003) Prediction of sudden sudden cardiac death: appraisal of the studies and methods assessing the risk of sudden arrhythmic death. *Circulation*, 108(1): 110–115.

KAYNAKLAR (devam ediyor)

- Ingle V K and Proakis J G** (2000) *Digital Signal Processing Using Matlab*. Pacific Grove, USA: Brooks/Cole Publishing Company.
- Izard C E, Simons R F, Haynes O M, Porges S W and Cohen B** (1991) Infant cardiac activity: Developmental changes and relations with attachment. *Dev Psychol*, 27(3): 432–439.
- İşler Y** (2009) *A Detailed Analysis of the Effects of Various Combinations of Heart Rate Variability Indices in Congestive Heart Failure*. Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü, Doktora tezi, İzmir.
- İşler Y, Avcu N, Kocaoğlu A ve Kuntalp M** (2008) Investigating the effects of the methods used in frequency domain measures in heart rate variability analysis. *In Proceedings of SIU2008, IEEE 16th Signal Processing and Communications Applications Conference*, Didim, Aydın, Turkey.
- İşler Y ve Kuntalp M** (2007a) Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure. *Comput Biol Med*, 37(10): 1502–1510.
- İşler Y ve Kuntalp M** (2007b) Investigating effects of wavelet entropy detailed measures in heart rate variability analysis. *In Proceedings of SIU2007, IEEE 15th Signal Processing and Communications Applications Conference*, Eskisehir, Turkey.
- İşler Y ve Kuntalp M** (2010) Heart rate normalization in the analysis of heart rate variability in congestive heart failure. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224(3): 453- 463.
- İşler Y ve Narin A** (2012) WEKA Yazılımında k-Ortalama Algoritması Kullanılarak Konjestif Kalp Yetmezliği Hastalarının Teşhisi, *Süleyman Demirel Üniversitesi Teknik Bilimler Dergisi*, 2(4): 21-29.
- İşler Y, Selver M A ve Kuntalp M** (2005) Effects of detrending in heart rate variability analysis. *In II. Mühendislik Bilimleri Genç Araştırmacılar Kongresi: MBGAK2005*, İstanbul, Turkey, 213–219.
- Jovic A and Bogunovic N** (2011) Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artif Intell Med*, 51: 175-186.
- Kamen P W** (1996) Heart rate variability. *Aust Fam Physician*, 25: 1087–1094.
- Kamen P W, Krum H and Tonkin A M** (1996) Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity. *Clin Sci*, 92: 201–208.
- Kamen P W and Tonkin AM** (1995) Application of the Poincare plot to heart rate variability: a new measure of functional status in heart failure. *Aust N Z J Med*, 25(1): 18–26.
- Kantelhardt J W, Koscielny-Bunde E, Rego H H A, Havlin S and Bunde A** (2001) Detecting long-range correlations with detrended fluctuation analysis. *Physica A*, 295: 441–454.

KAYNAKLAR (devam ediyor)

- Kleiger R E, Bigger J T, Bosner M S, Chung M K, Cook J R, Rolnitzky L M, Steinman R and Fleiss J L** (1991) Stability over time of variables measuring heart rate variability in normal subjects. *Am J Cardiol*, 68: 626.
- Kleiger R E, Miller J P, Bigger J T and Moss A J** (1987) Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am J Cardiol*, 59(4): 256–262.
- Kohler B U, Hennig C and Orglmeister R** (2002) The principles of software QRS detection: Reviewing and comparing algorithms for detecting this important ECG waveform. *IEEE Eng Med Biol*, 21(1): 42–57.
- Koza R J** (1992) *Genetic programming, on the programming of computers by means of natural selection*. Cambridge Massachusetts: MIT Press.
- Kreuzig E** (1993) *Advanced Engineering Mathematics*. Wiley, New York, 7th ed.
- Laguna P, Moody G B and Mark R G** (1998) Power spectral density of unevenly sampled data by least-square analysis: Performance and application to heart rate signals. *IEEE T Bio-Med Eng*, 45(6): 698–715.
- Landes R A, Scher M S, Sun M and Scwabassi R J** (1996) Characterization of heart rate dynamics in infants as a probe for neural state and age. *In 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1662–1663.
- Lanfranchi P A and Somers V K** (2002) Arterial baroreflex function and cardiovascular variability: interactions and implications. *Am J Physiol-Reg I*, 283: 815–826.
- Langley P, di Bernardo D, Allen J, Bowers E, Smith F E, Vecchietti S and Murray A** (2001) Can paroxysmal atrial fibrillation be predicted? *Comput Cardiol*, 28: 121–124.
- Leistner H L, Haddad G G, Epstein R A, Lai T L, Epstein M F and Mellins R B** (1980) Heart rate and heart rate variability during sleep in aborted sudden infant death syndrome. *J Pediatr*, 97(1): 51–55.
- Liao D, Barnes R W, Chambless L E and Heiss G** (1996) A computer algorithm to impute interrupted heart rate data for the spectral analysis of heart rate variability—the ARIC study. *Comput Biomed Res*, 29: 140–151.
- Liou D, Liou J and Liou C** (2013). *Learning Behaviors of Perceptron. Artificial Intelligence and Hybrid Systems*. iConcept Press.
- Lippman N, Stein K M and Lerman B B** (1993) Nonlinear predictive interpolation: a new method for the correction of ectopic beats for heart rate variability analysis. *J Electrocardiol*, 26: 14–19.

KAYNAKLAR (devam ediyor)

- Lippman N, Stein KM and Lerman B B** (1994) Comparison of methods for removal of ectopy in measurement of heart rate variability. *Am J Physiol*, 267(1): 411–418.
- Litvack D A, Oberlander T F, Carney L H and Saul J P** (1995) Time and frequency domain methods for heart rate variability analysis: a methodological comparison. *Psychophysiology*, 32: 492–504.
- Lomb N R** (1976) Least-squares frequency analysis of unequally spaced data. *Ap Sp Sc*, 39: 447–462.
- Lombardi F, Makikallio T H, Myerburg R J and Huikur H** (2001) Sudden cardiac death: role of heart rate variability to identify patients at risk. *Cardiovasc Res*, 50: 210–217.
- Luczak H and Lauring W J** (1973) An analysis of heart rate variability. *Ergonomics*, 16: 85–97.
- Lynn K S and Chiang H D** (2001) A two-stage solution algorithm for paroxysmal atrial fibrillation prediction. *Comput Cardiol*, 28: 405–407.
- Malik M and Camm A J** (1995) *Heart Rate Variability*. Futura Publishing Company.
- Malik M, Cripps T, Farrell T and Camm AJ** (1989) Prognostic value of heart rate variability after myocardial infarction. a comparison of different data-processing methods. *Med Biol Eng Comput*, 27: 603–611.
- Malik M, Farrell T, Cripps T and Camm A J** (1989) Heart rate variability in relation to prognosis after myocardial infarction: selection of optimal processing techniques. *Eur Heart J*, 10: 1060–1074.
- Malpas S C** (2002) Neural influences on cardiovascular variability: possibilities and pitfalls. *Am J Physiol-Heart C*, 282: 6–20.
- Marciano M L, Migax F, Acanfora D, Furgi G and Rengo F** (1994) Quantification of Poincare maps for the evaluation of heart rate variability. *Comput Cardiol*, 557–580.
- Martis R J and Chakraborty C** (2011) Arrhythmia disease diagnosis using neural network, svm, and genetic algorithm-optimized k-means clustering. *J Mech Med Biol*, 11(4): 897–915.
- Mateo J and Laguna P** (2003) Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE T Bio-Med Eng*, 50(3): 334–343.
- Mazursky J E, Birkett C L, Bedell K A, Ben-Haim S A and Segar J L** (1998) Development of baroreflex influences on heart rate variability in preterm infants. *Early Hum Dev*, 53(1): 37–52.
- McKinley R A, McIntire L K, Schmidt R, Repperger D W and Caldwell J A** (2011) Evaluation of eye metrics as a detector of fatigue. *Hum Factors*, 53 (4): 403–414.

KAYNAKLAR (devam ediyor)

- McNames J and Aboy M** (2006) Reliability and accuracy of heart rate variability metrics versus ECG segment duration. *Med Biol Eng Comput*, 44: 747–756.
- Mietus J E, Peng C K, Henry I, Goldsmith R L and Goldberger A L** (2002) The pNNx files: re-examining a widely used heart rate variability measure. *Heart*, 88: 378–380.
- Mitchell T M** (1997) *Machine Learning*. McGraw–Hill Companies Press.
- Mitov I P** (1998) A method for assessment and processing of biomedical signals containing trend and periodic components. *Medical Engineering and Physics*, 20(9): 660–668.
- Mokikallio T H, Seppanen T, Airaksinen J K E, Koistinen J, Tulppo M P, Peng C K, Goldberger AL and Huikuri H V** (1997) Dynamical analysis of heart rate may predict subsequent ventricular tachycardia after myocardial infarction. *Am J Cardiol*, 80: 779–783.
- Moody GB, Goldberger A L, McClennen S and Swiryn S** (2001) Predicting the onset of paroxysmal atrial fibrillation. *Comput Cardiol*, 28: 113–116.
- Mrowka R, Schluter B, Gerhardt D and Patzak A** (1996) Heart rate control in infants at high risk for sudden infant death syndrome SIDS. *Comput Cardiol*, 173–176.
- Myers GA, Martin G J, Magid N M, Barnett P S, Schaad J W, Weiss J S, Lesch M and Singer D H** (1986) Power spectral analysis of heart rate variability in sudden cardiac death: comparison to other methods. *IEEE T Bio-Med Eng*, 33(12): 1149–1156.
- Narin A ve İşler Y** (2012) Konjestif Kalp Yetmezliği Hastalarının Kalp Hızı Verileri Kullanılarak Teşhisi Üzerine Temel Bileşen Analizi, *IEEE 20th Signal Processing and Communications Applications Conference (SIU2012)*, Fethiye / Mugla / Turkey, 18-20 April.
- Narin A, İşler Y ve Özer M** (2013a) Konjestif Kalp Yetmezliğinin Kalp Hızı Değişkenliği Analizi ile Teşhisinde Sınıflandırıcı Başarımlarının Karşılaştırılması, *IEEE 21th Signal Processing and Communications Applications Conference (SIU2013)*, Girne / KKTC, 24-26 April.
- Narin A, İşler Y ve Özer M** (2013b) Investigating the Performance Improvement of HRV Indices in CHF using Statistical Feature Selection Methods, *Comput Biol Med*, under review.
- Nixon M S and Aguado A S** (2002) *Feature extraction and image processing*. Newness Butterworth-Heinemann, Woburn, 1st ed.
- Ori Z, Monir G, Weiss J, Sayhouni X and Singer D H** (1992) Heart rate variability frequency domain analysis. *Cardiol Clin*, 10(3): 499–533.
- Öztemel E** (2003) *Yapay Sinir Ağları*. Papatya Yayıncılık, İstanbul.

KAYNAKLAR (devam ediyor)

- Öztürk U** (2012) *Konjestif kalp yetmezliği hastalarının teşhisinde veri azaltma yöntemleri kullanılarak en yakın komşu sınıflandırıcı çalışma süresinin kısaltılması*. Bülent Ecevit Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Zonguldak.
- Pagani M** (2000) Heart rate variability and autonomic diabetic neuropathy. *Diabetes Nutr Metab*, 13(6): 341–346.
- Pagani M, Lombardi F, Guzzetti S, Rimoldi O, Furlan R, Pizzinelli P, Sandrone G, Malfatto G, Dell’Orto S, Piccaluga E, Turiel M, Baselli G, Cerutti S and Malliani A** (1986) Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympathovagal interaction in man and conscious dog. *Circ Res*, 59: 178–193.
- Pecchia L, Melillo P, Sansone M and Bracale M** (2011), Discrimination power of short-term heart rate variability measures for CHF assessment. *IEEE T Inf Technol B*, 15(1): 40–46.
- Penaz J, Roukenz J and Van der Waal H J** (1968) Spectral analysis of some spontaneous rhythms in the circulation. *In Leipzig, Germany: Biokybernetik, Karl Marx University*, 233–241.
- Peng C K, Buldyrev S V, Havlin S, Simons M, Stanley H E and Goldberger A L** (1994) Mosaic organization of DNA nucleotides. *Phys Rev E*, 49: 1685–1689.
- Pichot V, Gaspoz J M, Molltex S, Antoniadis A, Busso T, Roche F, Costes F, Quintin L, Lacour J R and Barthelemy J C** (1999) Wavelet transform to quantify heart rate variability and to assess its instantaneous changes. *J Appl Physiol*, 86(3): 1081–1091.
- Pikkujamsa S M, Makikallio T H, Sourander L B, Raiha I J, Puukka P, Skytta J, Peng C K, Goldberger A L and Huikuri H V** (1999) Cardiac interbeat interval dynamics from childhood to senescence: Comparison of conventional and new measures based on fractals and chaos theory. *Circulation*, 100: 393–399.
- Pincus S M** (1991) Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88: 2297–2301.
- Pincus S M and Viscarello R R** (1992) Approximate entropy: a regularity measure for fetal heart rate analysis. *Clin Obstet Gynecol*, 79: 249- 255.
- Pincus S M, Gladstone I M and Ehrenkranz R A** (1991) A regularity statistic for medical data analysis. *J Clin Monit Comput*, 7(4): 335–345.
- Pincus S M and Kalman E K** (2004) Irregularity, volatility, risk, and financial market time series. *Proc Natl Acad Sci*, 101 (38): 13709–13714.
- Pincus S M and Goldberger A L** (1994) Physiological time-series analysis: what does regularity quantify? *Am J Physiol*, 266 (4): 1643–1656.

KAYNAKLAR (devam ediyor)

- Pomeranz M, Macaulay R J B, Caudill M A, Kutz I, Adam D, Gordon D, Kilborn K M, Barger A C, Shannon D C, Cohen R J and Benson M** (1985) Assessment of autonomic function in humans by heart rate spectral analysis. *Am J Physiol*, 248:151–153.
- Porter G A J and Rivkees S A** (2004) Ontogeny of humoral heart rate regulation in the embryonic mouse. *Am J Physiol-Heart C*, 281: 401–407.
- Raymer M L, Punch W F, Goodman E D, Kuhn L A and Jain A K** (2000) Dimensionality reduction using genetic algorithms. *IEEE T Evolut Comput*, 4(2): 164–171.
- Rentero N, Cividjian A, Trevaks D, Pequignot J M, Quintin L and McAllen R M** (2002) Activity patterns of cardiac vagal motoneurons in rat nucleus ambiguus. *Am J Physiol-Reg I*, 283: 1327–1334.
- Richman J S and Randall M J** (2000) Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol-Heart C*, 278: 2039- 2049.
- Riera A R P, Ferreira C, Filho C F** (2008) The enigmatic sixth wave of the electrocardiogram: the U wave. *Cardiol J*, 15(5): 408- 21.
- Rosenblatt F** (1957) The Perceptron--a perceiving and recognizing automaton. *Report*, 85-460-1, Cornell Aeronautical Laboratory.
- Rosenstock E G, Cassuto Y and Zmora E** (1999) Heart rate variability in the neonate and infant: Analytical methods, physiological and clinical observations. *Acta Paediatr*, 88: 477–482.
- Rosso O A, Blanco S, Yordanova J, Kolev V, Figliola A and Başar E** (2001) Wavelet entropy: a new tool for analysis of short duration brain electrical signals. *J Neurosci Methods*, 105: 65–75.
- Rosso O A, Quian Q R, Blanco S, Figliola A and Başar E** (1998) Wavelet–entropy applied to brain signal analysis. *In Proceedings of the IX European Signal Processing Conference, EUSIPCO–98*, 4: 2445–2448.
- Rottman J N, Steinman R C, Albrecht P, Bigger J T J, Rolnitzky L M and Fleiss J L** (1990) Efficient estimation of the heart period power spectrum suitable for physiologic or pharmacologic studies. *Am J Cardiol*, 66: 1522–1523.
- Sayers B M** (1973) Analysis of heart rate variability. *Ergonomics*, 16: 17–32.
- Sabeti M** (2009) Entropy and complexity measures for EEG signal classification of schizophrenic and control participants. *Artif Intell Med*, 47 (3): 263–274.
- Scargle J D** (1982) Studies in astronomical time series analysis. II. statistical aspects of spectral analysis of unevenly spaced data. *Astrophys J*, 263: 835–853.

KAYNAKLAR (devam ediyor)

- Schechtman V L, Henslee J A and Harper R M** (1998) Developmental patterns of heart rate variability in infants with persistent apnea of infancy. *Early Hum Dev*, 50: 251–262.
- Schilling R J and Sandra L H** (1999) *Applied Numerical Methods for Engineers*. Pacific Grove, CA, USA: Brooks/Cole Publishing Company.
- Shamsham F and Mitchell J** (2000) Essentials of the diagnosis of heart failure. *Am Fam Physician*, 61: 1319–1328.
- Shannon C E** (1948) A mathematical theory of communication. *AT&T Tech J*, 27: 379–423 and 623–656.
- Shimazu M, Ojima S, Takasugi S I, Nejime T and Kamishima T** (1999) Time–frequency analysis of heart rate variability using complex discrete wavelet transform. In *Annual International Conference of the IEEE Engineering in Medicine and Biology – Proceedings*, 958.
- Shlens J** (2009) A Tutorial on Principal Component Analysis, <http://www.snl.salk.edu/~shlens/pca.pdf>
- Smith A L, Reynolds K J and Owen H** (2007) Correlated Poincare indices for measuring heart rate variability. *Australas Phys Eng Sci Med*, 30(4):336–341.
- Spicer C C and Lawrence C J** (1987) Statistical analysis of heart rates in subsequent victims of sudden infant death syndrome. *Stat Med*, 6: 159–166.
- Stein P K, Le Q and Domitrovich P P** (2008) Development of more erratic heart rate patterns is associated with mortality post-MI. *J Electrocardiol*, 41(2): 110–115.
- Taqqu M S, Teverovsky V and Willinger W** (1995) Estimators for long range dependence: An empirical study. *Fractals*, 3: 785–798.
- Tarvainen M P, Ranta-aho P O and Karjalainen P A** (2002) An advanced detrending method with application to HRV analysis. *IEEE T Bio-Med Eng*, 49(2): 172–175.
- Task Force [of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology]** (1996) Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur Heart J*, 17: 354–381.
- Thakre T P and Smith M L** (2006) Loss of lag-response curvilinearity of indices of heart rate variability in congestive heart failure. *BMC Cardiovasc Disord*, 6, 27.
- Torrence C and Compo G P** (1998) A practical guide to wavelet analysis. *B Am Meteorol Soc*, 79(1): 61–78.
- Valafar F** (2001) Applications of Neural Networks in Medicine and Biological Sciences. In *Intelligent Control Systems Using Soft Computing Methodologies*. CRC.

KAYNAKLAR (devam ediyor)

- Van Eck H J R, Kors J A and Van Herpen G** (2005) The U wave in the electrocardiogram: A solution for a 100-year-old riddle. *Cardiovasc Res*, 67: 256-262.
- Vatanabe S** (1985) *Pattern recognition: human and mechanical*. New York: JohnWiley and Sons Inc.
- Vapnik V** (1998) *Statistical Learning Theory*. New York: John Wiley and Sons.
- Vornanen M, Ryokkynen A and Nurmi A** (2002) Temperature-dependent expression of sarcolemmal K⁺ currents in rainbow trout atrial and ventricular myocytes. *Am J Physiol-Reg I*, 282: 1191–1199.
- Weber E J M, Molenaar C M and van der Molen M W** (1992) A nonstationarity test for the spectral analysis of physiological time series with an application to respiratory sinus arrhythmia. *Psychophysiology*, 29(1):55–65.
- Webster J G** (1993) *Design of cardiac pacemakers*. TAB–IEEE Press.
- Webster J G** (1998) *Medical instrumentation–application and design*. Houghton Mifflin Company, 3. baskı.
- Wiklund U, Akay M and Niklasson U** (1997) Short-term analysis of heart-rate variability by adapted wavelet transforms. *IEEE Eng Med Biol*, 16(5): 113–118 and 138.
- Wilbur J and James P** (2005) Diagnosis and management of heart failure in the outpatient setting. *Prim Care*, 32: 1115–1129.
- Williams T D, Chambers J B, Henderson R P, Rashotte M E and Overton J M** (2002) Cardiovascular responses to caloric restriction and thermo-neutrality in C57BL/6J mice. *Am J Physiol-Reg I*, 282: 1459–1467.
- Wolf M M, Varigos G A, Hunt D and Sloman J G** (1978) Sinus arrhythmia in acute myocardial infarction. *Med J Aust*, 2: 52–53.
- Woo M A, Stevenson W G, Moser D K, Trelease R B and Harper R H** (1992) Patterns of beat-to-beat heart rate variability in advanced heart failure. *Am Heart J*, 123:704–710.
- Yeragani V K, Sobolewski E, Jampala V C, Kay J, Yeragani S and Igel G** (1998) Fractal dimension and approximate entropy of heart period and heart rate: Awake versus sleep differences and methodological issues. *Clin Sci*, 95: 295–301.
- Yu S N and Lee M Y** (2012a) Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability. *Comput Biol Med*, 42: 816- 825.
- Yu S N and Lee M Y** (2012b) Conditional mutual information-based feature selection for congestive heart failure recognition using heart rate variability. *Comput Methods Programs Biomed*, 108: 299- 309.

KAYNAKLAR (devam ediyor)

Yuan Q (2011) Epileptic EEG classification based on extreme learning machine and nonlinear features. *Epilepsy Res*, 96 (1- 2): 29–38.

Yun K (2012) Decreased cortical complexity in methamphetamine abusers. *Psychiatry Res*, 201 (3): 226–32.

Zambroski C H (2003) Qualitative analysis of living with heart failure. *Heart Lung*, 32(1): 32–40.

ÖZGEÇMİŞ

Ali NARİN Manisa’da doğdu; ilk ve orta öğrenimini Manisa’da tamamladı. 2006 yılında girdiği Süleyman Demirel Üniversitesi Elektronik ve Haberleşme Mühendisliği bölümünden 2011 yılında mezun oldu. 2011 yılında Bülent Ecevit Üniversitesi (Zonguldak Karaelmas Üniversitesi) Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim dalında yüksek lisansa başladı. Halen BEÜN Elektrik Elektronik Mühendisliği Anabilim Dalı’nda araştırma görevlisi olarak çalışmaktadır.

ADRES BİLGİLERİ

Adres : B.E.Ü. Elektrik-Elektronik Mühendisliği
İncivez – ZONGULDAK

Tel : 0372 257 40 10 – (1341)

E-posta : alinarin45@gmail.com