

Demografik Özelliklerin Koroner Arter Hastalığına Etkisinin Analizi

Kemal AKYOL¹, Şafak BAYIR², Baha ŞEN³

¹ Karabük Üniversitesi, Bilgisayar Mühendisliği Bölümü, Karabük

² Karabük Üniversitesi, Bilgisayar Mühendisliği Bölümü, Karabük

³ Yıldırım Beyazıt Üniversitesi, Bilgisayar Mühendisliği Bölümü, Ankara

kemalakyol48@gmail.com, safakbayir@gmail.com, bsen@ybu.edu.tr

Özet: Koroner arterler kalbin kendisini besleyen damarlardır. Koroner arter hastalığı bu damarlarda meydana gelen tıkanıklıklarla ortaya çıkar ve çok yaygın bir hastalıktır. Bu hastalığa neden olan bir çok risk faktörünün olduğu bilinmektedir. Bu çalışmamızda sınıflandırma teknikleri ile modifiye edilebilen risk faktörlerinin koroner arter hastalığına etkisi analiz edilmiştir. Çalışmamızda kullanılan verisetleri biyokimya ve hemogram laboratuvar test sonuçlarını içermektedir. Analizler ile elde edilen başarımlar ölçütleri net bir şekilde göstermektedir ki, hipertansiyon ve diyabet mellitus faktörlerinin KAH üzerinde etkisi vardır. Ayrıca, bu çalışma koroner arter hastalığı gibi medikal verilerin işlenmesi ve sınıflandırılmasında random forest ve destek vektör makineleri algoritmalarının kullanılabileceğini bize göstermiştir.

Anahtar Sözcükler: Koroner Arter Hastalığı, Değiştirilebilir Risk Faktörleri, Bilgi Keşfi, Sınıflandırma.

Analysis of Impact of Demographic Characteristics on Coronary Artery Disease

Abstract: Coronary arteries are vessels supplying the heart itself. Coronary artery disease occurs with occlusions occurring in the veins and very common disease. It is known that there are many risk factors that cause disease. In this study, effect of modifiable risk factors on coronary artery disease is analyzed by classification techniques. The dataset used in this study include the biochemistry and hemogram laboratory test results. Performance metrics obtained by analysis clearly showed that hypertension and diabetes mellitus factors have an effect on coronary heart disease. In addition, this study showed us that can be used random forests and support vector machines algorithms in processing and classification of medical data such as coronary artery disease.

Keywords: Coronary Artery Disease, Modify Risk Factors, Knowledge Discovery, Classification.

1. Giriş

Koroner Arter Hastalığı(KAH) kalp adalesini besleyen ve koroner arterler olarak adlandırılan atar damarların daralma veya tıkanması ile kan akımının biraz veya tamamen kesilmesine bağlı olarak ortaya çıkan hastalıklardır. Bu hastalığın nedeni halk arasında damar sertliği olarak bilinen aterosklerozdur ve en önemli özelliği ileri evrelerinde sağlamları ciddi derecede tehliyeke sokan kalp krizine yol açabilmesidir [1]. Bu hastalığı tetikleyen lipid risk faktörlerinin yanı sıra nonlipid risk faktörleri de bulunmaktadır. Bu anlamda bilinen ve bilinmeyen bir çok risk faktörünün olduğu bilinmektedir. Ve bu alanda bir çok çalışma yapılmıştır. Ulusal Kolesterol Eğitim Programı'nın 2001'de yayınlanan III. Yetişkin tedavi panelinde koroner arter hastalığı ile ilgili nonlipid risk faktörlerinden modifiye edilebilecek olan yani kontrol altında tutulabilecek olan risk faktörleri çalışmamızın başlangıç noktasını oluşturmaktadır ve bu faktörler aşağıda verilmiştir [2]:

- Hipertansiyon

- Sigara içiyor olmak
- Diyabetes Mellitus
- Fazla kiloluluk/Obezite
- Fiziksel inaktivite
- Aterojenik diyet
- Trombojenik/ hemostatik durum

Üzerinde çalıştığımız veriler ile yüksek lisans tez çalışmasında[3] "hangi test değerleri KAH için önemlidir?" sorusuna yanıt aranmıştır. Bu çalışmada ise "hangi test değerleri önemlidir?" sorusuna yanıt aramaktan ziyade KAH için önemli olan test değişkenleri ile hipertansiyon ve diyabetik mellitus demografik özelliklerinin KAH üzerindeki etkisi araştırılmıştır. Tabii ki, literatürde görüleceği üzere, KAH'nı doğrudan veya dolaylı olarak etkileyen bir çok faktör vardır. Yani, hastanın tüm değerleri normal iken bile KAH olma ihtimali söz konusudur ki bu durum gizli faktörlerin etkisi olarak da açıklanabilir. Doğru sınıflandırma oranları ve başarımlar ölçütleri çerçevesinde modifiye edilebilen risk faktörlerin-

den hipertansiyon ve diyabet mellitus tanısının olduğu verisetleri ile çalışılmıştır. Verisetindeki verilerin az olması nedeniyle diğer risk faktörleri ile çalışılamamıştır.

Bu çalışmanın içeriği şu şekildedir: 2. kısımda KAH ile ilgili yapılmış çalışmalara yer verilmiştir. 3. kısımda bilgi keşfinin önemi sunulmuştur. 4. kısımda sınıflandırma hakkında bilgi verilmiştir. 5. kısımda analiz için geliştirilen uygulama ele alınmıştır. 6. kısımda verilerin analizi ile elde edilen sonuçlar verilmiştir. Son bölümde sonuçlar ve öneriler sunulmuştur.

2. İlgili Çalışmalar

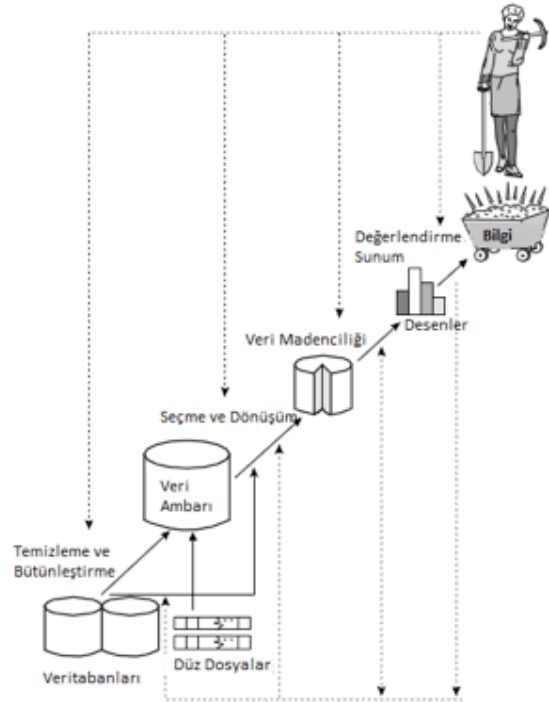
KAH ile ilgili olarak yapılmış bir çok çalışma bulunmaktadır. Bu çalışmalardan bazıları şunlardır: Framingham çalışmasında bilgi keşfi neticesinde elde edilen belirli kriterler çerçevesinde 10 yıl içinde KAH olma riskini analiz etmişlerdir [4]. Bir başka çalışma olan PROCAM çalışmasında KAH'nı önleme amaçlı bir sistem geliştirilmiştir. Bu sistem ile erken tanı ve hastalığı mümkün olduğunca önleyebilme çalışmaları yapılmıştır [5]. TEKHARF çalışmasında özellikle yetişkin Türk bireylerinin KAH riskini skorlama çalışması yapılmıştır. Bu çalışma bireylere ait 10 yıllık bilgi birikimi ile gerçekleştirilmiştir [6]. Wang ve arkadaşları, Framingham fonksiyonunu kullanarak geleneksel risk faktörlerinin KAH üzerinde farklı derecelerde etkili olduğu ve diğer risk faktörlerinin de bu hastalığa ayrıca katkıda bulunduğuna dair önemli bir sonuca varmışlardır [7]. Karaolis ve arkadaşları Apriori algoritması tabanlı veri madenciliği sistemi geliştirmişlerdir. Bu sistem, KAH ile alakalı risk faktörlerini tanımlamak için kullanılmıştır. Bu çalışmaya göre, sigara KAH'nı tüm durumlarda doğrudan etkileyen önemli bir risk faktörüdür [8]. Srinivas ve arkadaşları, karar ağacı, yapay sinir ağları ve bayes sınıflayıcıları gibi veri madenciliği tekniklerini kullanarak etkili bir şekilde kalp krizini tahmin etmeye çalışmışlardır [9]. Abdullah ve Rajalaxmi, KAH'nın tahmininde doktorlara yardımcı olması için random forests algoritmasını kullanarak veri madenciliği tekniği geliştirmişlerdir [10]. Lojistik risk skoru isimli bir çalışmada, KAH riski tahmin edilmektedir. Bu hastalık ile ilgili risk faktörlerinin tanımlanması çerçevesinde akut koroner sendromların ve KAH olan hastaların tedavisinde önemli bir yere sahiptir [11]. Bir başka çalışmada, biyokimya ve hemogram laboratuvar test verileri Türk Kardiyoloji Derneği(TKD) bilgileri referans alınarak lojistik regresyon yöntemi ile test edilmiştir. KAH için önemli bulunan test değişkenleri lojistik regresyon modeline dahil edilerek KAH analizi gerçekleştirilmiştir. Doktorların KAH ris-

kini değerlendirebilmeleri için yeni veri girişine imkan veren bir sistem tasarlamışlardır [12].

3. Bilgi Keşfi

Bilgi keşfi, veriden anlamlı bilgiyi elde etmek için gerekli olan yöntem ve teknikleri içerir. Günümüzde, büyük veri setlerinden anlamlı bilgiyi çıkarmak için iş dünyası, bankacılık ve tıbbi alan gibi çok çeşitli alanlarda hızla artan bir kullanımı mevcuttur. Özellikle tıp alanındaki verilerin önem derecesi bu alandaki uygulamaları daha da önemli kılmaktadır. Şekil 1'de gösterilen bilgi keşif sürecinin adımları şu şekildedir [13]:

1. Veri temizleme: Gürültülü ve anlamsız verileri verisetinden çıkarmak,
2. Veri bütünleştirme: Gerekli olduğu durumda birden çok veri kaynağını birleştirebilmek,
3. Veri seçme: Analiz için gerekli olan verileri belirlemek,
4. Veri dönüşümü: Verinin kullanılabilir hale dönüşümünü gerçekleştirmek,
5. Veri madenciliği: Veri örüntülerini sınıflandırabilmek için gerekli yöntem ve teknikleri uygulamak,
6. Örüntü değerlendirme: Başarı ölçüm kriterleri çerçevesinde performans başarı analizi yapmak
7. Bilgi sunumu: Elde edilen anlamlı bilginin kullanıcıya sunumunu gerçekleştirmek.



Şekil 1. Bilgi Keşif Süreci Adımları [13]

4. Sınıflandırma

Sınıflandırma, yaygın olarak kullanılan ve analizlerde önemli bir yeri olan bir makine öğrenme tekniğidir. Girdi olarak verilen bilginin çeşitli özelliklere ve niteliklere göre bir sınıflandırma algoritması tarafından sınıflara atanması sürecidir. Girdi, her biri bir sınıf etiketi ile etiketlenecek gözlem veya örneklerden oluşan bir eğitim kümesidir. Çıktı ise öğrenen sistemin verilen girdiye dayalı olarak atadığı sınıf etiketidir. Elde edilen sonuçlar, başarılı sınıflandırma oranları ve performans ölçüm değerleri sınıflandırmanın başarısını bize sunar [14]. Özellik tanımlama, sınıflandırma sürecinin önemli bir adımıdır ve aslında problemin çözümünün çekirdeğini oluşturur. Bu anlamda, KAH ile ilgili olarak laboratuvar test sonuçlarının analizi için Türk Kardiyoloji Derneği(TKD) verileri esas alınmıştır.

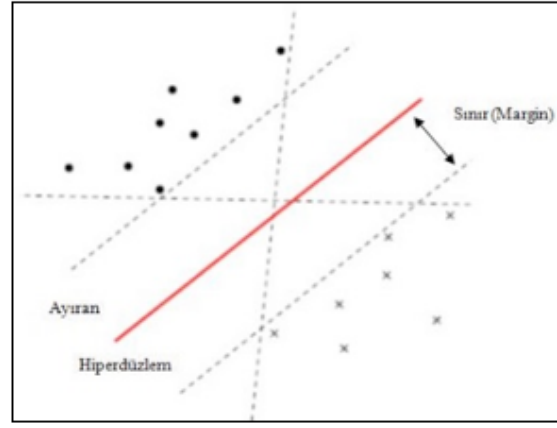
4.1. Random Forest(RF)

Random forest(Rastgele orman) algoritması Breiman tarafından tanıtılan çok popüler ve etkili bir makina öğrenme tekniğidir [15]. RF yöntemi çok sayıda karar ağacı ile oluşturulan orman sınıflandırıcısıdır ve bu yöntemle sınıflama veya regresyon ağaçları kurulabilmekte ve kümeleme yapılabilmektedir. Veri setindeki “sınıf değişkeni” kategorik ise sınıflama, sürekli ise regresyon ağaçları kurulmaktadır [16].

RF, tüm değişkenler arasından en iyi dalı kullanarak her bir düğümü dallara ayırmak yerine, her bir düğümde rastgele bir biçimde seçilen değişkenler arasından en iyisini kullanarak her bir düğümü dallara ayırır. Her veri seti orijinal veri setinden yer değiştirmeli olarak üretilir. Sonra rastgele özellik seçimi kullanılarak ağaçlar geliştirilir [17].

4.2. Destek Vektör Makineleri (DVM)

DVM sınıflandırması, iki ayrı sınıfa ait verilerin en iyi şekilde birbirinden ayrılması için en uygun ayırıcı aşırı düzlemi kullanır. Bu işlem, iki sınıf uzayının birbirinden olabildiğince uzakta temsil edilmesiyle gerçekleştirilir. İki sınıf arasında karar aralığı sınırları(margin) bulunmaktadır. Bu sınırlar birbirinden ne kadar uzak olursa bulunan düzlem o kadar uygun olur. Sınırlar üzerindeki veri setine ait örneklerle destek vektörler denir. Şekil 2’de görüldüğü üzere, bir verisetini ayırmak için pek çok hiper düzlem(şekil üzerindeki kesikli çizgiler) olabilir ancak en iyi genellemeyi yapan yani en iyi sınır(margin) olan kırmızı çizgidir [18].



Şekil 2. İki sınıflı bir problemde DVM hiper düzlemleri [18]

Bu çalışmada kullandığımız bu iki sınıflandırma algoritması ile laboratuvar test bilgileri verilen bir veri setine ait sınıf etiketlerini öğrenebilen veya tahmin edebilen bir model oluşturulması amaçlanmıştır.

5. Uygulama

5.1. Veriler

Geriye dönük ve olgu-kontrol araştırması şeklinde gerçekleştirilen bu çalışmanın verileri, 01.01.2011-11.10.2011 tarihleri arasında Yıldırım Beyazıt Üniversitesi Ankara Atatürk Eğitim ve Araştırma Hastanesi Kardiyoloji ve diğer servislerinden temin edilmiştir. Çalışma verilerinden koroner kalp hastasına ait veriler, KAH grubuna, diğer veriler ise kontrol grubuna dahil edilmiştir.

ICD-10 Uluslararası Hastalık Kodları kitapçığı referans alınarak KAH tanısı veriseti oluşturulmuştur. Olgularda KAH ile ilgili olarak biyokimya ve hemogram laboratuvar test değerlerine ilişkin referans değerlerin tanımlanmasında TKD’nin 2002’de yayınladığı “Koroner Kalp Hastalığı Korunma ve Tedavi Klavuzu” esas alınmıştır. Referans değer, iki değer arasında olabileceği gibi belirli bir eşik değerinin altında veya üstünde de olabilir. Bu klavuzda belirtilen ve araştırmamızda kullanılan koroner arter hastalığı risk değerleri şunlardır [19]:

- Erkeklerde 45 ve daha üzeri, kadınlarda 55 ve daha üzeri yaşında olma,
- Serum LDL kolesterol değerinin ≥ 130 mg/dl olması ve/veya total kolesterolün ≥ 200 mg/dl olması,
- Serum trigliserid düzeylerinin ≥ 150 mg/dl olması,
- Serum HDL kolesterol düzeylerinin < 40 mg/dl olması.

5.2. Önişleme

Çalışma için temin edilen veriler geliştirilen yazılım ile sisteme yüklenmiştir. Veri setindeki biyokimya ve hemogram laboratuvar test değişkenlerinden KOLESTEROL, TRIGLİSERİD, HDL ve LDL değişkenleri ve ayrıca CINSİYET ve YAS belirleyici faktör değişkenleri modele dahil edilen değişkenlerdir. KAH değişkeni sonuç değişkeni olmak üzere veri setinin düzenlenmesi sağlanmıştır. Kategorik değişken verileri text formatında bir içeriğe sahiptir ve bu çalışma için bu verilerin hepsi kategorik olarak 1 ve 0 dönüşümü sağlanmıştır. Örneğin KAH tanısı pozitif ise 1, negatif ise 0 dönüşümü sağlanmıştır. Diğer yandan, TKD verileri doğrultusunda, yaş ve cinsiyet verilerinin ikisini birden ele alan bir dönüşüm yapılmıştır. Şöyle ki; cinsiyeti erkek ve yaşı 44 ve üzeri ise veya cinsiyeti kadın ve yaşı 54 ve üzeri ise 1 diğer durumlarda 0 kategorik dönüşümü sağlanmıştır. Sürekli değişken verileri normalize edilerek veri setleri oluşturulmuştur. Bir test değişkenine ait referans değer aralığı TKD verileri doğrultusunda sisteme tanıtılmış ve bu doğrultuda verisetleri elde edilmiştir. Tablo 1’de önişlem aşamasından geçirilmeden önce verisetinde mevcut olan değişkenler listesi görülmektedir.

Değişken ismi	Veri türü	Açıklama
Cinsiyet	Kategorik	Cinsiyet
Yas	Sürekli	Hastanın yaşı
Hastalık Tanısı	Kategorik	KAH Tanısı
LDLKolesterol	Sürekli	LDL kolesterol değeri
HDLKolesterol	Sürekli	HDL kolesterol değeri
Trigliserid	Sürekli	Trigliserid değeri
Kolesterol	Sürekli	Kolesterol değeri
Hipertansiyon	Kategorik	Hipertansiyon tanısı
Diyabet Mellitus	Kategorik	Diyabet Mellitus tanısı

Tablo 1. Verisetindeki değişkenler listesi

HT (Veriseti 1)		DM (Veriseti 2)		HT ve DM (Veriseti 3)	
0	1	0	1	0	1
30	30	16	16	46	46

Tablo 2. Oluşturulan verisetleri

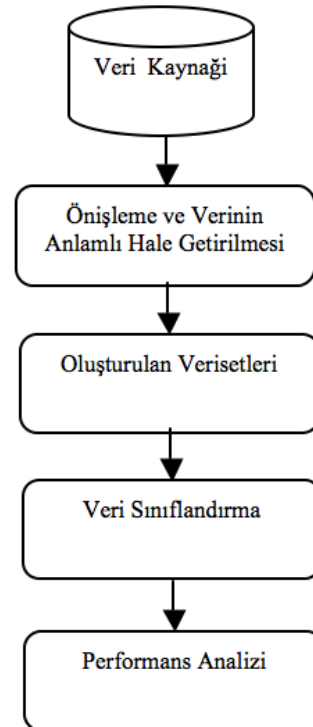
		Veri seti 1	Veri seti 2	Veri seti 3
KAH	0	19	10	29
	1	41	22	63
Eğitim		30	16	46
Test		30	16	46

Tablo 3. KAH tanısı ve sınıflandırma için kullanılacak eğitim ve test veri sayısı bilgileri

Analizler için oluşturulan her bir veri setinde 2 kategorik ve 4 sürekli değişken bulunmaktadır. Bu verisetlerinin %50’lik kısmı eğitim için, diğer %50’lik kısmı ise modelin testi için ayrılmıştır. Eğitim ve test verilerinin yarısı demografik özelliğinden oluşmaktadır. Ancak her bir verisetindeki KAH hastalığı tanısı olan hasta sayısı farklılık gösterebilmektedir. Tablo 2 ve 3’deki veriler incelendiğinde görülmektedir ki, örneğin 30 hastada HT tanısı var, diğer 30 hastada HT tanısı yok iken bu verisetindeki KAH sayısı 41, KAH tanısının olmadığı hasta sayısı 19’dur.

5.3. Geliştirilen Uygulama

Bu çalışmada biyokimya ve hemogram laboratuvar test değerleri ile sınıflandırma yapılmıştır. Bu verisetlerinin oluşturulması sonrasında RF ve DVM sınıflandırma algoritmaları ile hipertansiyon ve diyabet mellitus faktörlerinin KAH üzerindeki etkisi araştırılmıştır.



Şekil 3'deki iş akışında görüldüğü üzere, her bir veriseti üzerinde gerçekleştirilen analizlerin işlem adımları şu şekildedir:

- Laboratuvar ve tanı bilgileri yüklenir ve gürültülerden arındırılmış olarak veriler hazırlanır.
- Anlamlı bilgiler elde edilir ve verisetleri oluşturulur.
- Belirlenen kurgu doğrultusunda sınıflandırma işlemi gerçekleştirilir.
- Oluşturulan modellerin başarısı performans değerlendirme ölçütleri çerçevesinde test edilir.

6. Analizler ve Elde Edilen Sonuçlar

Üzerinde çalıştığımız veriler ile ilgili bilgiler ve performans başarı analizleri Tablo 4 ve 5'te

görüldüğü üzere tablo formatında sunulmuştur. Uygulanan sınıflandırma algoritmaları neticesinde veri setindeki KAH tanısı verileri ile analiz sonucu elde edilen değerler karşılaştırılarak “Doğru Sınıflandırma Oranı(DSO)” hesaplanmıştır. Bu tabloda, KAH'na ait “Beklenen” ve “Gözlenen” değerler görülmektedir. Örneğin, RF algoritması ile hipertansiyonun KAH'na etkisi incelendiğinde görülmektedir ki, koroner arter hastası olmayan 2 kişi yanlış sınıflandırılarak koroner arter hastası grubuna, koroner arter hastası olan 5 kişi de yanlış sınıflandırılarak koroner arter hastası olmayan gruba atanmıştır. Başka bir ifadeyle koroner arter hastası olmayanların %50,0'si, koroner arter hastası olanların %90,0'ı doğru tahmin edilmiştir.

Verisetleri		Random Forest			Destek Vektör Makineleri		
		Hayır	Evet	Başarı(%)	Hayır	Evet	Başarı(%)
Hipertansiyon tanılı olgular ve diğer olgular ile oluşturulan veriseti ile KAH analizi	No	5	5	50,00	9	1	90,00
	Yes	2	18	90,00	3	17	85,00
	Genel Başarı			76,66	Genel Başarı		86,66
Diyabet Mellitus tanılı olgular ve diğer olgular ile oluşturulan veriseti ile KAH analizi	No	3	0	100,00	3	0	100,00
	Yes	5	8	61,53	5	8	61,53
	Genel Başarı			68,75	Genel Başarı		68,75
Hipertansiyon ve Diyabet Mellitus tanılı olgular ve diğer olgular ile oluşturulan veriseti ile KAH analizi	No	12	3	80,00	12	3	80,00
	Yes	5	26	83,87	4	27	87,09
	Genel Başarı			82,60	Genel Başarı		84,78

Tablo 4. KAH duyarlılığı oluşturan faktörler ile elde edilen tahmin sonuçları

Sınıflandırma algoritmalarının performansını ölçmek için duyarlılık(sensitivity,SN), seçicilik(specificity,SP) ve gerçek doğru oranı(TPR) ölçütleri kullanılmıştır. Bu değerlendirme ölçütleri sırasıyla Eşitlik 1'de verilmiştir [20].

$$SN = TP / TP + FN, SP = TN / TN + FP, TPR = TP / TP + FP \quad (1)$$

Eşitlik 1'de yer alan başarı ölçütleri hesaplama-ları 4 temel değere dayandırılarak hesaplanır. Bunlar;

- Doğru Pozitif (DP) :KAH tanısı olan hastaların doğru bir şekilde KAH olarak sınıflandırılması,
- Yanlış Pozitif (YP) :KAH tanısı olmayan hastaların yanlış bir şekilde KAH olarak sınıflandırılması,
- Yanlış Negatif (YN) :KAH tanısı olan hastaların yanlış bir şekilde KAH tanısı yok olarak sınıflandırılması,
- Doğru Negatif (DN) :KAH tanısı olmayan hastaların doğru bir şekilde KAH tanısı yok olarak sınıflandırılması.

Sınıflandırıcılar	Verisetleri	SN	SP	TPR	DSO
Random Forest	Veriseti 1	71,42	78,26	50,00	76,66
	Veriseti 2	37,50	100,00	100,00	68,75
	Veriseti 3	70,58	89,65	80,00	82,60
Destek Vektör Makinaları	Veriseti 1	75,00	94,44	90,00	86,66
	Veriseti 2	37,50	100,00	100,00	68,75
	Veriseti 3	75,00	90,00	80,00	84,78

Tablo 5. Sınıflandırma değerlendirme sonuçları

7. Sonuçlar ve Öneriler

KAH, sağlıklı bireylerde %5-10 sıklıkla görülürken, laboratuvar test değerlerinin referans aralığı değerlerini aşan hastalarda %80'e ulaşabilmektedir. Ayrıca yaş ile birlikte KAH olma riski artmaktadır.

Sınıflandırma ile veriden anlamlı bilgi elde edilerek bir bakıma bilgi keşfi icra edilmiş olur. Sınıflandırma, eğitim ve gözlem yolu ile verinin çeşitli sınıflardan hangisine ait olduğunu bulan önemli bir makine öğrenme tekniğidir.

Geriye dönük ve olgu-kontrol araştırması şeklinde 01.05.2011-28.11.2011 tarihleri arasında Kardiyoloji ve diğer servislerdeki 92 adet hasta verisi üzerinde yapılan bu çalışmada rutin biyokimya ve hemogram laboratuvar test değerleri analiz edilerek modifiye edilebilen risk faktörlerinden hipertansiyon ve diyabetik mellitus tanılarının KAH' na etkileri araştırılmıştır.

Çalışma verileri, hasta gurubu ve kontrol grubundan oluşmaktadır. KAH verileri kardiyoloji bölümündeki uzman doktorlar tarafından KAH tanısı ile işleme alınmış verilerden, kontrol grubu hastalarına ait veriler ise kardiyoloji bölümü haricindeki diğer bölümlerdeki hastalardan oluşmaktadır. Kontrol grubu verileri her ne kadar kardiyoloji bölümüne ait verilerden oluşmamış olsa da hastanın kalp hastası olabileceği ihtimali göz önünde bulundurulmalıdır. Elde edilen sonuçlar, RF ve DVM sınıflandırma algoritmalarının başarılı bir şekilde sınıflandırma yaptığını ve bu sınıflandırma algoritmalarının KAH tanısı için alanında uzman yetkili kişilere yardımcı olabileceğini bize göstermiştir.

Ayrıca elde edilen veri setindeki olgu sayısının az olması da göz önünde bulundurulması gereken diğer önemli bir ayrıntıdır. Veri setindeki olgu sayısı ne kadar çok olursa veri uzayı o kadar iyi tanımlanacaktır ki böylece algoritmaların başarı performansı da o kadar iyi olacaktır. Verisetindeki verilerin az olması nedeniyle bu çalışmamızda diğer risk faktörleri ile çalışılamamıştır. Sonraki çalışmalarımızda modifiye edilebilen tüm risk faktörlerini içeren bir çalışma yapılması planlanmaktadır.

8. Kaynaklar

- [1] İnternet: http://www.tkdcd.org/public/uploads/files/pdf/saglikli_yasam/koroner_arter_hastaliklari.pdf (2014).
- [2] Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on De-

tection Evaluation, and Treatment of High Blood Cholesterol in Adults Final Report, National Cholesterol Education Program National Heart, Lung, and Blood Institute. National Institutes of Health, NIH Publication No. 02-5215, 2:14-15, (2002).

- [3] Akyol, K. "Biyokimya ve Hemogram Laboratuvar Test Sonuçlarının Lojistik Regresyon Yöntemiyle Analizi", Karabük Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Karabük, (2012).

- [4] Kannel, W. B., McGee, D. and Gordon, T., "A general cardiovascular risk profile: the Framingham Study", Am J Cardiol, 38:1-1, (1976).

- [5] Assmann, G., Cullen, P. and Schulte, H., "Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the Prospective Cardiovascular Münster(PROCAM) study", Circulation, 105(3):310-5, (2002).

- [6] Onat, A., Keleş, İ., Çetinkaya, A., Başar, Ö., Yıldırım, B., Erer, B., Ceyhan, K., Eryonucu, B. ve Sansoy, V., "Prevalence of Coronary Mortality and Morbidity in the Turkish Adult Risk Factor Study: 10-year Follow-up Suggests Coronary Epidemic", Türk Kardiyol Arş, 29:8-19, (2001).

- [7] Wang, Z., Hoy and W.E., "Is the Framingham coronary heart disease absolute risk function applicable to Aboriginal people?", Med.J.Australia, 182(2): 66-69, (2005).

- [8] Karaolis, M., Moutiris, J.A. and Pattichs, L., "Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining With Decision Trees", IEEE Transactions on IT in Biomedicine, 14(3):559-66, (2010).

- [9] Srinivas, K., Rao, G.R. and Govardhan, A., "Analysis of Coronary Heart Disease and Prediction of Heart Attack in Coal Mining Regions Using Data Mining Techniques", 5th IntConf on Computer Science and Education, 1344-1349, (2010).

- [10] Abdullah, A.S. and Rajalaxmi, R.R., "A Data mining Model for predicting the Coronary Heart Disease using Random Forest Classifier", International Conference on Recent Trends in Computational Methods, Communication and Controls (ICON3C 2012), 23, (2012).

- [11] Chambless, L.E., Dobson, A.J. and Patterson, C.C. and Raines B., "On the use of a logistic risk score in predicting risk of coronary heart disease", Stat Med, 9(4):385-96, (1990).

- [12] Akyol, K., Çalık, E. ve Şen, B., "Biyokimya ve Hemogram Laboratuvar Test Sonuçlarının Lojistik Regresyon Yöntemiyle Analizi", XIV. Akademik Bilişim, 350-354, Uşak, (2012).

- [13] Han, J. and Kamber, M., "Data Mining:

- Concepts and Techniques”, Morgan Kaufmann Publishers, San Francisco, 5-7 (2006).
- [14] Emel, G.G. ve Taşkın, Ç., “Veri Madenciliğinde Karar Ağaçları ve Bir Satış Analizi Uygulaması”, Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 6(2):224, (2005).
- [15] Breiman, L., “Random forests. Machine Learning”, 45(1): 18, (2001).
- [16] Akman, M., Genç, Y. ve Ankaralı, H., “Random Forests Yöntemi ve Sağlık Alanında Bir Uygulama”, Türkiye Klinikleri J Biostat, 3(1): 37, (2011).
- [17] Akar, Ö. ve Güngör, O., “Rastgele orman algoritması kullanılarak çok bantlı görüntülerin sınıflandırılması”, Jeodezi ve Jeoinformasyon Dergisi, 1(2):139-146, (2012).
- [18] Pala, T., “Tıbbi Karar Destek Sisteminin Veri Madenciliği Yöntemleriyle Gerçekleştirilmesi”, Marmara Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 20, İstanbul, (2013).
- [19] İnternet: <http://old.tkd.org.tr/kilavuz/k11.htm>, TKD Koroner Kalp Hastalığı Korunma ve Tedavi Kılavuzu (2014).
- [20] Tripathi, S., Singh, K.K., Singh, B.K. and Mehrotra, A., “Automatic Detection of Exudates in Retinal Fundus Images using Differential Morphological Profile”, International Journal of Engineering and Technology (IJET), 5(3): 2028-2029, (2013).