



# Demografik Özelliklerin Koroner Arter Hastalığına Etkisinin Analizi

Kemal AKYOL, Şafak BAYIR, Baha ŞEN





# İçindekiler

- ◆ Giriş
- ◆ Bilgi Keşfi
- ◆ Sınıflandırma
- ◆ Uygulama
- ◆ Sonuç ve Öneriler



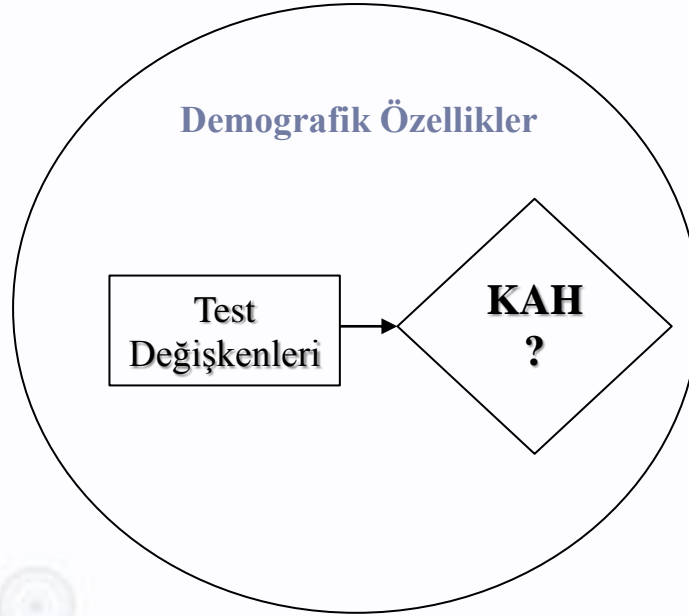
# Giriş

- ◆ **Koroner Arter Hastalığı(KAH)** : Koroner arterler kalbi besleyen damarlardır. Bu damarlarda meydana gelen tıkanıklıklarla ortaya çıkan yaygın bir hastalıktır.
- ◆ **Koroner arter hastalığı ile ilgili nonlipid risk faktörleri:**
  - Hipertansiyon +
  - Sigara içiyor olmak
  - Diyabetes Mellitus +
  - Fazla kiloluluk/Obezite
  - Fiziksel inaktivite
  - Aterojenik diyet
  - Trombojenik/ hemostatik durum



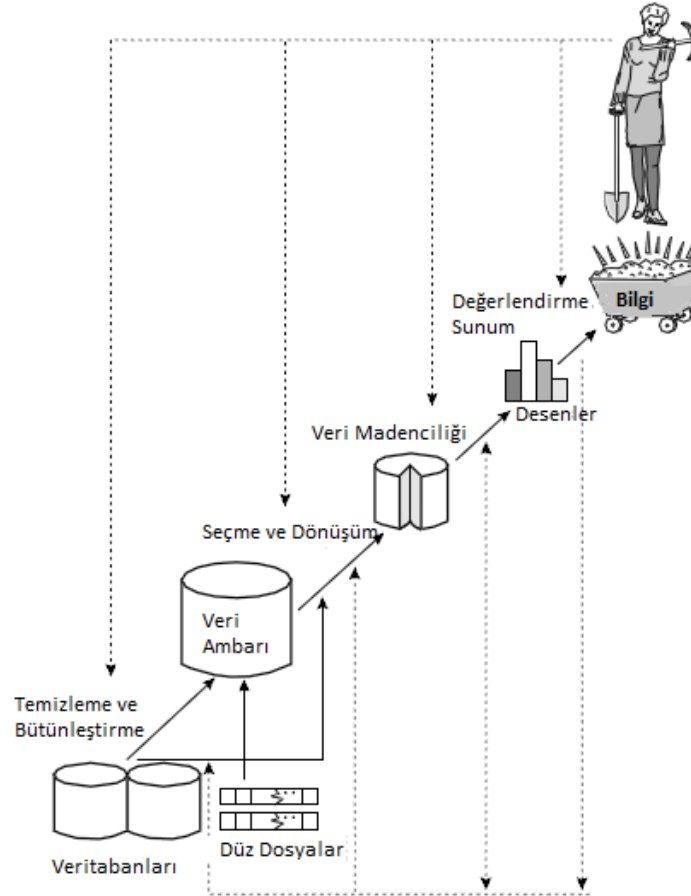
# Amaç

- ◆ Hipertansiyon ve diyabetik mellitus demografik özelliklerinin KAH üzerindeki etkisi araştırılmıştır.



# Bilgi Keşfi

- ◆ Bilgi keşfi, veriden anlamlı bilgiyi elde etmek için gerekli olan yöntem ve teknikleri içerir.



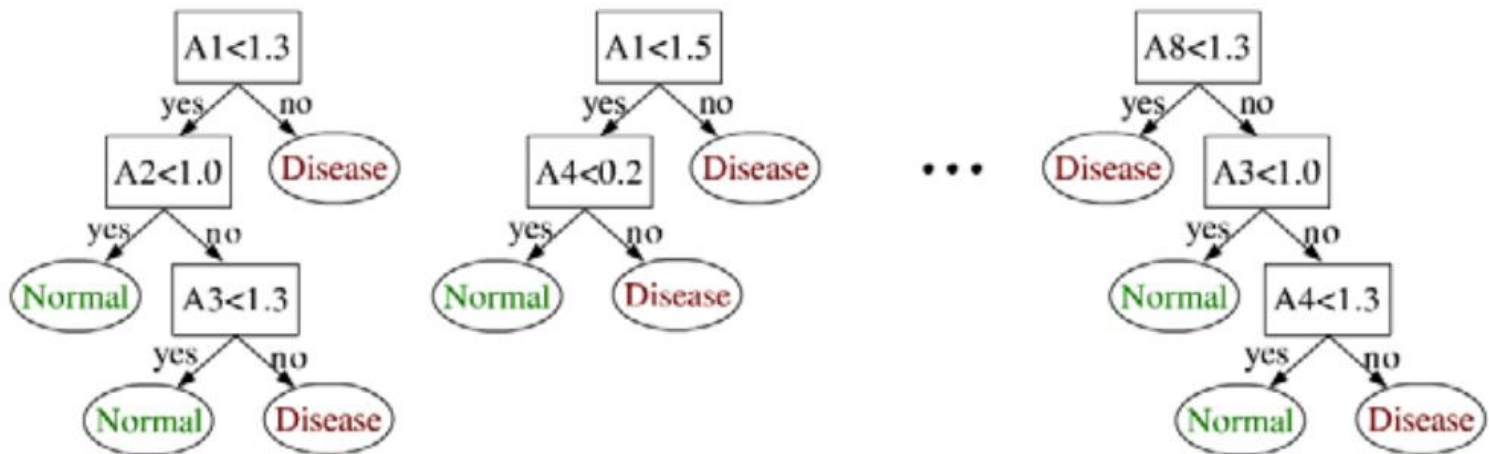
# Sınıflandırma

- ◆ Girdi olarak verilen bilginin çeşitli özelliklere ve niteliklere göre bir sınıflandırma algoritması tarafından sınıflara atanması sürecidir.



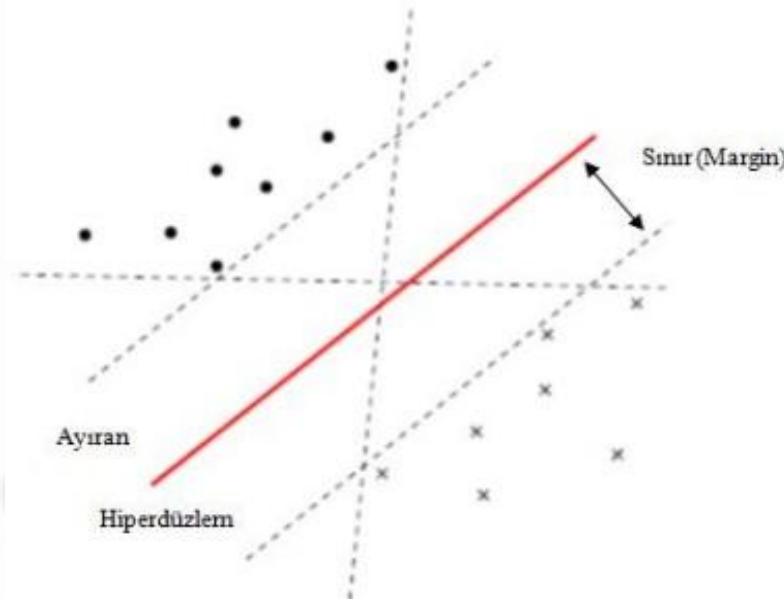
# Random Forest Sınıflandırıcısı

- ◆ RF yöntemi çok sayıda karar ağacı ile oluşturulan bir sınıflandırıcıdır.
- ◆ 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.



# 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.



# Uygulama

- ◆ Veriler: 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.



# Uygulama

- ◆ ICD-10 Uluslararası Hastalık Kodları referans alınarak KAH tanısı veriseti oluşturulmuştur.

<b>I25</b>	<b>Kronik iskemik kalp hastalığı</b>
	<i>Hariç:</i> kardiyovasküler hastalık BŞT (I51.6)
<b>I25.0</b>	<b>Aterosklerotik kardiyovasküler hastalık</b>
<b>I25.1</b>	<b>Aterosklerotik kalp hastalığı</b>
	Koroner (arter) ateromu
	Koroner (arter) aterosklerozi
	Koroner (arter) hastalığı
	Koroner (arter) sklerozi
<b>I25.2</b>	<b>Eski miyokard enfarktüsü</b>
	İyileşmiş miyokardiyal enfarktüs
	EKG veya diğer özel çalışmalarla tanı konmuş fakat semptom vermeyen geçirilmiş miyokardiyal enfarktüs
<b>I25.3</b>	<b>Kalp anevrizması</b>
	Anevrizma, mural
	Anevrizma, ventriküler
<b>I25.4</b>	<b>Koroner arter anevrizması</b>
	Koroner arteriyovenöz fistül, kazanılmış
	<i>Hariç:</i> konjenital koroner (arter) anevrizması (Q24.5)
<b>I25.5</b>	<b>İskemik kardiyopati</b>
<b>I25.6</b>	<b>Sessiz miyokardiyal iskemi</b>
<b>I25.8</b>	<b>Kronik iskemik kalp hastalığının diğer formları</b>



# Uygulama

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



# Uygulama - Önışleme

- ◆ Kategorik deęişken verileri text formatındadır ve bu alıřma iin kategorik olarak 1 ve 0 dnüşümü saęlanmıřtır. Örneęin KAH tanısı pozitif ise 1, negatif ise 0 dnüşümü saęlanmıřtır.
- ◆ TKD doęrultusunda, yař ve cinsiyet verileri: cinsiyeti erkek ve yařı 44 ve üzeri ise veya cinsiyeti kadın ve yařı 54 ve üzeri ise 1 dięer durumlarda 0 kategorik dnüşümü saęlanmıřtır.
- ◆ Sürekli deęişken verileri normalize edilerek veri setleri oluřturulmuřtur.



# Uygulama

ALT	AST	GLUKD	HCT	HDL	HGB	KOLEST	KREAT	LDL	LVIM	LVMY	MCH	MCHC	MCV	MPV	PCT	POW	PLT	POTAĐ	SEC	SDW	NA	TRIGLİ	URE	Wbc	Hastal	Cinsiy	Yas	Hasta	HT	HL	DM	Sigara	AO
26	19	80	45,3	33,4	155,5	189	1,1	126,8	2,3	30,4	31,8	34,2	93	8,2	0,245	16,4	239	4,2	4,87	13,1	144	144	23	7,6	2E+09	E	31	Evet	Hayır	Hayır	Hayır	Evet	
26	16	106	42,5	25,5	142,5	166	0,9	110,3	2,7	22,8	30,5	34,1	89,3	7,3	0,236	16,6	324	4,4	4,76	13	143	151	23	11,7	2E+09	E	53	Evet	Hayır	Hayır	Hayır	Hayır	
40	23	102	46,2	32,1	168,8	175	1,1	84,9	3,4	30,2	31,2	36,4	85,6	8,7	0,198	16,5	229	4,8	5,4	13,3	139	290	30	11,4	2E+09	E	30	Hayır	Hayır	Hayır	Hayır	Hayır	
26	22	98	46,5	61,9	164	194	1,1	107,5	1,7	22,1	31,6	35,3	89,5	8	0,169	16	212	4,6	5,19	14,8	139	123	43	7,7	2E+09	E	55	Hayır	Hayır	Hayır	Hayır	Hayır	
25	17	85	40,9	42,7	139,9	233	0,8	157,3	2,9	46,9	30,4	33,9	89,5	7,2	0,186	16,5	237	4,8	4,57	13,9	141	165	20	6,3	2E+09	K	56	Hayır	Hayır	Hayır	Hayır	Hayır	
9	16	123	33,1	51,3	117,7	134	0,8	71,9	1,5	30,7	31,3	35,3	88,6	9	0,156	16,6	173	4,4	3,73	14,2	145	54	54	4,8	2E+09	K	75	Evet	Evet	Evet	Evet	Hayır	
18	19	110	35,7	40	121,1	161	1,2	93,2	1,6	21,7	29,7	33,9	87,6	9,6	0,311	16,2	325	5	4,07	13,6	139	139	42	7,5	2E+09	K	67	Hayır	Evet	Evet	Evet	Hayır	
25	30	183	30,7	38,5	104,4	115	1,2	58,5	0,7	3,2	30,5	33,7	90,7	10,8	0,104	16	97	4,9	3,39	15,1	136	90	81	20,9	2E+09	K	62	Evet	Evet	Evet	Evet	Hayır	
43	27	93	45,2	39,1	151,1	195	1,2	116,5	1,9	36,2	31	33,5	92,7	8,5	0,152	16,9	179	4,3	4,88	12,7	140	197	16	5,3	2E+09	E	41	Hayır	Hayır	Hayır	Hayır	Hayır	
22	23	100	39,6	72,4	142,2	238	0,8	144,8	1,9	30,2	31,2	35,8	87,1	7,9	0,228	15,8	290	4,5	4,54	13,8	138	104	26	6,3	2E+09	K	68	Hayır	Hayır	Hayır	Hayır	Hayır	
19	23	103	36,1	44,8	123,3	158	1,2	77	2,1	32	26,1	34,1	76,3	9,2	0,189	16,4	207	4,5	4,73	16,5	141	181	31	6,4	2E+09	E	56	Evet	Hayır	Hayır	Hayır	Hayır	
19	16	186	37,8	46,3	129,7	177	2	107,9	2,8	34,1	29,5	34,3	85,9	9,5	0,126	17,5	133	5	4,39	14,4	139	114	61	8,3	2E+09	E	69	Evet	Hayır	Evet	Evet	Hayır	
16	22	79	36,3	52,5	133,3	193	0,8	124,1	2	29,3	32,2	36,6	88,2	9,8	0,205	16	210	4,3	4,12	13,1	139	82	17	6,8	2E+09	K	39	Evet	Hayır	Hayır	Hayır	Hayır	
8	15	86	35,6	29,1	106,6	160	0,8	96,1	2,4	20,8	19,7	29,8	66,3	8,8	0,353	17,5	633	5,2	3,37	21,1	140	164	26	11,3	2E+09	E	38	Evet	Hayır	Hayır	Hayır	Hayır	
14	20	84	34,1	40,7	132,2	171	1,7	107,6	1,4	21,7	21,8	24	64,3	7,1	0,151	16,6	187	4,2	4,08	14,4	140	141	38	4	2E+09	E	63	Hayır	Hayır	Hayır	Hayır	Hayır	

HDL	KOLESTER	LDL	TRIGLİSER	Cinsiyet	Yas	Hastalık_Tanisi	HT	HL	DM	Sigara	AO
33,4	189	126,8	144	E	31	Evet	Hayır	Evet	Hayır	Hayır	Evet
25,5	166	110,3	151	E	53	Evet	Hayır	Hayır	Hayır	Hayır	Hayır
32,1	175	84,9	290	E	30	Hayır	Hayır	Hayır	Hayır	Hayır	Hayır
61,9	194	107,5	123	E	55	Hayır	Hayır	Hayır	Hayır	Hayır	Hayır
42,7	233	157,3	165	K	56	Hayır	Hayır	Hayır	Hayır	Hayır	Hayır
51,3	134	71,9	54	K	75	Evet	Evet	Evet	Evet	Hayır	Hayır
40	161	93,2	139	K	67	Hayır	Evet	Evet	Evet	Hayır	Hayır
38,5	115	58,5	90	K	62	Evet	Evet	Hayır	Evet	Hayır	Hayır
39,1	195	116,5	197	E	41	Hayır	Hayır	Hayır	Hayır	Hayır	Hayır



# Uygulama

Veri Seti İşlemleri Model Oluşturma İşlemleri İstatistik Bilgiler Yeni Veri Analizi Çıkış

C:\Users\Kaan\Desktop\veriSeti.xlsx

Birim Adı : TÖM BİRİMLER (\*)

	KOLESTROL	TRİGLİSERİD	HDLKOLESTEROL	LDLKOLE
▶	163	81	54	93
	145	30	74	65
	302	199	38	244
	259	165	25	141
	267	187	47	153
	128	44	57	62
	185	139	42	115
	218	180	52	142
	253	192	43	192
	181	63	67	101
	261	186	21	133
	199	195	36	132
	249	123	62	115
	127	114	46	58
	290	121	41	145
	227	256	51	145
	268	144	26	102

Veri Seti İşlemleri

**TÖM BİRİMLER (\*) TESTLERİ**

Seçim Yapılan Test Sayısı: 9

İşlem Detayını Göster

Sonuç Bilgisini Göster

Veri Setini Düzenle

Tümünü Seç / İptal

BİLİRUBİNDİREK

KOLESTROL

TRİGLİSERİD

HDLKOLESTEROL

LDLKOLESTEROL

WBC

**İŞLEM SONUCU**

Toplam 3211 adet kayıt silindi...

Null Değer olan kayıt yok

Veri Setindeki Analiz Edilebilir Kayıt Sayısı=169

**VERİ ÖNİŞLEME İŞLEMLERİ - ADIM 4**

**İŞLEM:VERİ SETİNİN İŞLENMESİ**

Toplam Kayıt Sayısı: 169

İşlem Durumu:169/169

Genel İşlem Durumu:4/4



# Uygulama – Oluşturulan Veri Setleri

## ◆ Oluşturulan Verisetleri

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

		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



# Uygulama – Oluşturulan Veri Setleri

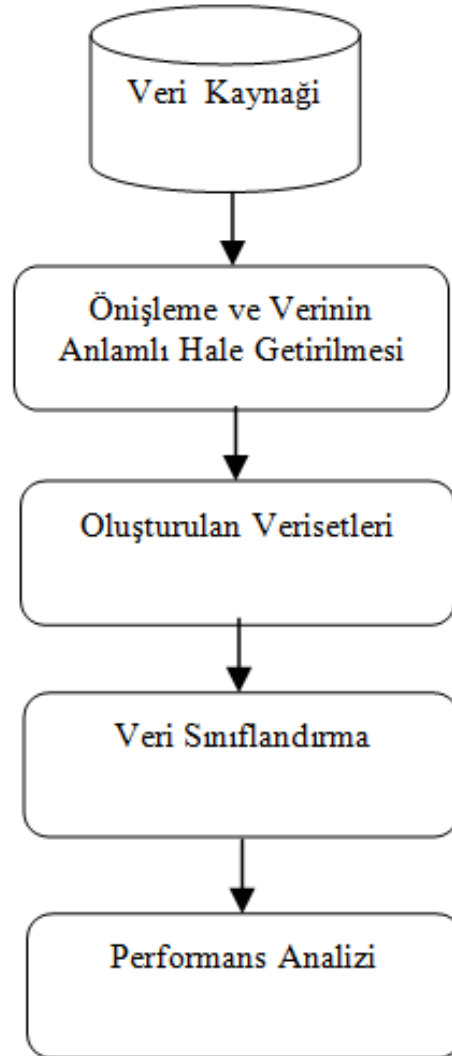
## ◆ Oluşturulan Verisetleri

HT (Veriseti 1)		DM (Veriseti 2)		HT ve DM (Veriseti 3)	
0	1	0	1	0	1
<b><u>30</u></b>	<b><u>30</u></b>	16	16	46	46

		Veri seti 1	Veri seti 2	Veri seti 3
KAH	0	<b><u>19</u></b>	10	29
	1	<b><u>41</u></b>	22	63
Eğitim		<b><u>30</u></b>	16	46
Test		<b><u>30</u></b>	16	46



# Geliştirilen Uygulama



# Analizler ve Elde Edilen Sonular

## ◆ Deęerlendirme lutleri

Duyarlılık(sensitivity-SN), Seicilik(specificity-SP) ve gerek doęru oranı(TPR) lutleri...

$$SN = DP/DP+YN,$$

$$SP = DN/DN+YP, \quad TPR = DP/DP+YP$$

- Doęru Pozitif (DP)
- Yanlıř Pozitif (YP)
- Yanlıř Negatif (YN)
- Doęru Negatif (DN)



# Analizler ve Elde Edilen Sonuçlar

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 5.** Sınıflandırma değerlendirme sonuçları

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

# 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.
- ◆ 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.



# Sonuçlar ve Öneriler

- ◆ 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.



# Kaynaklar

- ◆ **İnternet:**  
[http://www.tkdcd.org/public/uploads/files/pdf/saglikli\\_yasam/koroner\\_arter\\_hastaliklari.pdf](http://www.tkdcd.org/public/uploads/files/pdf/saglikli_yasam/koroner_arter_hastaliklari.pdf) (2014).
- ◆ **Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection 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).**
- ◆ **Akyol, K. "Biyokimya ve Hemogram Laboratuvar Test Sonuçlarının Lojistik Regresyon Yöntemiyle Analizi", Karabük Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Karabük, (2012).**
- ◆ **Kannel, W. B., McGee, D. and Gordon, T., "A general cardiovascular risk profile: the Framingham Study", Am J Cardiol, 38:1-1, (1976).**
- ◆ **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).**
- ◆ **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).**
- ◆ **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).**
- ◆ **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).**
- ◆ **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).**

**Katılımınız için teşekkürler...**

