

ABPM Ölçümü Olmaksızın Karar Ağaçları Algoritması ile Non-Dipper/Dipper Öngörüsü

Aysun ALTIKARDEŞ^a, Hasan Erdal^b, Fevzi BABA^c, Ali Serdar Fak^d

^a Bilgisayar Teknolojileri Bölümü, Marmara Üniversitesi, İstanbul
^{b, c} Elektrik-Elektronik Mühendisliği Bölümü, Marmara Üniversitesi, İstanbul
^d İç Hastalıkları AD, Marmara Üniversitesi, İstanbul

Non-Dipper/Dipper Prediction without ABPM by Decision Trees Algorithm

Abstract: The aim of this study was to design an expert system to predict the dipping or non-dipping pattern by using several basic clinical and laboratory data without Ambulatory Blood Pressure Monitoring (ABPM) through an artificial intelligence algorithm. Data Mining is a technique which extracts information from data sets by using a combination of both statistical analysis methods and artificial intelligence algorithms. Also in this study, it is used.

Sixty-five patients (mean age 51 ± 7 years, 40 females,) were included in the study. Systolic and diastolic dipping was found in 13 and 15 % of the patients, respectively. The data sets were tested using the J48 decision tree-inducing algorithm. In this classification algorithm was implemented on 56 instances and the number of attributes was reduced from 47 to 21 by filtering. The training data set was 66% of instances. So, 19 instances splitted for test. When this algorithm was run, the instances were predicted non-dipper/dipper correctly in 73.6842 %. 0.02 seconds were to build model. Sensitivity = 70% and Specificity = 78%

This pilot study shows that a machine learning algorithm can help in the prediction of diurnal blood pressure pattern relying on some basic demographic, clinical and laboratory data, with a reasonable accuracy.

Key Words: Decision Tree; Diabetes; ABPM; Classification

Özet: Marmara Üniversitesi Hastanesi Hipertansiyon ve Ateroskleroz Ünitesi ile birlikte, ABPM ölçümü olmaksızın sadece klinik muayene ve laboratuvar tetkikleri ile elde edilen hasta parametrelerine dayanılarak, Non-dipper özelliğın ön görölüp göröllemeyeceğine ışık tutmaya yönelik ortak bir çalışmadır. Diabetes Mellitus (DM) ve Hipertansif (HT) hastalarda zaten çoğu zaman yapılması gereken klinik ve laboratuvar incelemelerine ait sonuçlarla, yapay zekâ algoritmaları ve istatistiksel analiz metotlarının birlikte kullanıldığı veri madenciliği tekniğiyle dipper/non-dipper hastaları öngörmeye yönelik bir uzman sistem tasarımı amaçlanmaktadır.

Bu çalışmada 65 gönüllü hasta (ortalama yaş 51 ± 7 , 40 kadın) yer almıştır. Bu hastalarda %13 sistolik, %15 diastolik dipping özelliği bulunmuştur. Bu deneyde 56 denegin %66'sı eğitim için ayrılmıştır. Böylece, J48 karar ağacı algoritması 19 denek üzerinde test edilmiştir. Nitelik sayısı, non-invazif yöntem olan 24 saatlik kan basıncı takip cihazı ABPM ile elde edilen gece/gündüz, sistolik ve diastolik kan basınçlarına ait 9 adet nitelik kullanılmaksızın 47'den 21'e indirgenmiş

ve %73.6842 doğru sınıflandırma başarısı, %70 duyarlılık ve % 78 belirlilik ortaya konmuştur.

Bu pilot çalışma, bir makine öğrenme algoritmasının makul bir hassasiyetle, bazı temel demografik, klinik ve laboratuvar verilerine dayanarak günlük kan basıncı seyrinin öngörülmesinde yardımcı olabileceğini göstermiştir.

Anahtar Kelimeler: Karar Ağacı; Diyabet; ABPM; Sınıflandırma

1. Giriş

DM, ciddi metabolik hastalıklar arasında prevalansı nüfus artışına, yaşlanmaya, kentleşmeye, fiziksel aktiviteye ve obezite prevalansının yükselmesine bağlı olarak tüm dünyada giderek artan epidemik bir hastalıktır. Sinsi seyirli bir hastalık olduğundan prevalansının saptanması güçlük yaratmaktadır. Hemen hemen tüm toplumlarda görülmesine karşın diyabet prevalansı ırka bağlı olarak anlamlı farklılıklar göstermektedir [1, 2].

Diyabetes Mellitus, dünya genelinde ve tüm yaş gurplarında WHO'nun yaptığı hesaplara göre, 2000 yılında prevalansı % 2.8 iken 2030 yılında % 4.4 olarak hesap edilmiştir. 2000 yılında; dünyada toplam 171 milyon olan diyabetli hasta sayısı, 2003 yılı itibarı ile 194 milyon olmuş ve 2030 yılında 366 milyona ulaşması beklenmektedir[3,4,2-5]. Diyabetli hasta sayının yaklaşık olarak % 97'si tip 2 DM sınıflamasına girmektedir[6,7]. Dünya'da tip 2 diyabet prevalansının önümüzdeki on yıl içinde yaklaşık olarak %40 artarak 150 milyondan 210 milyona ulaşması beklenmektedir [8].

Ülkemizde diyabet sıklığını belirlemeye yönelik 1997–1998 yıllarında gerçekleştirilen ve rastgele olarak seçilmiş 20 yaş ve üzerindeki 24,788 kişiyi kapsayan ‘Türkiye Diyabet Epidemiyoloji (TURDEP) Çalışması’nın sonuçlarına göre, ülkemizde tip 2 diyabet prevalansı % 7.2, IGT sıklığı ise % 6.7’dir. Bu oranlara dayanarak 2000 yılı nüfus sayısına göre, ülkemizde 2,6 milyonun üzerinde diyabetli ve 1,8 milyon civarında IGT’li kişinin yaşadığı sanılmaktadır [9,10,11,12].

2006 yılında Anadolu Kardiyoloji dergisinde yayınlanan orijinal araştırma sonucuna göre Türkiye’de 305bin olan DM insidansı hızla artmakta, prevalansı da benzer şekilde yükselmektedir [13].

Sağlıklı bireylerde gündüz KB ortalama değerine göre, gece uykusu sırasında ortalama KB değeri %10 ve üzeri şeklinde düşüş göstermektedir. Bu duruma “Dipper Paterni” denilmektedir. Bu fizyolojik durumun bozularak gece uykusu esnasında ortalama KB değerinin düşmemesi ise “Non-Dipper Paterni” olarak adlandırılmaktadır[14, 15].

Bireylerin yaşam kaliteleri üzerinde büyük etkisi olan uyku ile sağlık arasında çift yönlü bir ilişki vardır ve bireyin sağlıklı olarak fonksiyonlarını yerine getirebilmesi açısından önemlidir. Huzursuz geçen uyku dönemi tıbbi ve psikiyatrik morbidite ve tüm nedenlere bağlı ölümlerdeki artışın yanı sıra kötü sağlık sonuçları (inflamasyon, insülin direnci, metabolik sendrom, non-dipper paterni) için risk faktörleri ile bağlantılıdır. Yapılan klinik ve epidemolojik çalışmalar hipertansif hastalarda”non-dipper” özelliğın KVVH riskini bağımsız olarak daha da arttırdığını göstermektedir. Ayrıca DM’lu hastalarda da hipertansiyondan bağımsız olarak önemli oranlarda “non-dipper” özellik bulunmektedir.

Bugün dipper/nondipper KB özelliğı klinik olarak ancak 24 saatlik KB kaydı ile

ayırt edilir. Non-invazif bir yöntem olmakla birlikte, hastanın 24 saat bir kayıt cihazını taşıması hastanın yaşam kalitesi, zaman, emek ve maliyet açısından türlü olumsuzluklar içerir. Yapılan bu pilot çalışma ile ABPM cihazı verileri olmaksızın doktor teşhisine en yakın şekilde non-dipper öngörüsü yapabilen bir sistem tasarlanarak bu olumsuzlukların giderilmesi esas hedefdir.

2. Gereç ve Yöntem

Önerilen sistemde örneklem kümeleri üzerinde yapay zekâ algoritmaları ve istatistiksel analiz yöntemleri bileşimi uygulanarak bilgi çıkarılması tekniği olan veri madenciliği kullanılmıştır.

Makine öğrenimi - Machine Learning (ML); mevcut veri kümelerinin analiziyle bilgisayarların “öğrenme” tekniklerini geliştiren, yapay zekânın bir dalını temsil etmektedir. Bu şekilde, geçmiş tecrübelerinden bilgi çıkararak, yeni bir uyarana karşı uygun tepki verebilmeleri mümkün olabilmektedir[16]. Bu bağlamda, Yeni Zelanda Waikato Üniversitesi tarafından Java programlama dilinde geliştirilerek, ticari amaç gütmeyen (GNU General Public License) kullanıma sunulmuş olan ve programcının kendi kodlamasına olanak tanıyan Weka 3.7.4 Developer versiyonu bu çalışmada tercih edilmiştir[17,18].

Bu aşamada denek grubunu oluşturan 40’ı kadın olmak üzere, tip 2 diabetli ve normotansif toplam 65 gönüllü hastanın verileri weka ortamına aktarılmıştır. Veri toplandıktan sonraki aşama anlamlı yollarla bilgiyi organize etmektir. Bu bağlamda analize başlanmıştır.

İlk olarak veri toplama sürecinde, bazı deneklerden bazı verilerin alınamamasından dolayı oluşan null veriler tespit edilmiştir. Çalışmanın, hedef organ hasarı ve dolayısıyla KVH ve sonuçlarını öngörmeye olabilecek etkileri göz önünde bulundurularak, null verisi olan deneklerin tamamı örneklem kümesinden çıkarılmış, böylece instance yani denek sayısı 56’ya indirgenmiştir.

Veri tablolarında denekler (instances) satırlarda, nitelikler (attributes) ise sütunlarda yer almaktadırlar.

Tıp alanında önde gelen literatür çalışmaları ışığında ve öznelik seçimi için fikir veren algoritmaların sonuçları göz önüne alınarak, her yapılan deneyde değerlendirme dışında bırakılan nitelikler, “Nitelik Sayısı Tablosu”nda yeşil ile kapatılmıştır (Tablo 1). Böylece nitelik sayısı 21’e indirgenmiştir.

Karar ağaçları güvenilir ve etkin karar vermeye yaygın şekilde yardımcıdırlar. Doğruluk oranlarının yüksekliği ve ortaya çıkarılan bilginin kolay anlaşılmasını sağlayan grafik gösterimlerinden dolayı sınıflandırmada tercih edilmektedirler [19-20]. Karar ağaçları özel bir uygulanabilirlik alanı olmamakla birlikte, tıp biliminin alt alanlarında, sağlık hizmetleri uygulamalarında [19] ve karar verme seviyelerinde geniş kullanım alanına sahiptirler. Çeşitli hastalardan veri toplamak tıbbi bilgilerin çıkarımına ve karar destek sistemine yardımcı olabilmektedirler.

Endüktif öğrenme algoritmalarından biri olan karar ağaçları, verilerin/örneklerin eğitimde kullanılmasıyla sınıflandırma ağacının oluşmasını sağlamaktadır. Temeli böl ve yönet (divide and conquer) stratejisine dayanmaktadır [21, 22]. Bu yöntemin avantajı, anlaşılabilir kurallara dönüştürülebilirleridir [23]. Karar ağaçları temel olarak tıbbi veriler için teşhis modelleri oluşturmakta kullanılmaktadır [24,25]. Tıbbi verilerden kalıplar/pattern keşfedilmeye çalışıldığında [26], yetersiz kaldıkları ortaya çıkmaktadır.

J. Ross Quinlan tarafından 1986 yılında Sydney Üniversitesi’nde öne sürülmüş C4.5 karar ağaçları, ID3 (Iterative Dichotomiser 3) eşleştirme öğrenme algoritmasına

(matching learning algorithm) dayanmakta ve örneklerden öğrenmenin en iyi bilinen aracı kabul edilmektedir [20, 27, 28]. C4.5 örnek edinme / kıyaslama ile öğrenme sistemidir (benchmark learning system) ve sınıflama performansı en iyi olanlardan birisidir [29]. Buradaki kıyaslama veya başkalarından öğrenme: en iyi yöntemlerin, uygulamaların ve örneklerin sürekli olarak araştırılması anlamında dinamik bir kavramdır. Temelde ID3 algoritmasının gelişmiş versiyonu 1993 yılında aynı araştırmacı tarafından yayınlanmıştır [30,31]. Ayrıca çok yaygın olarak kullanılan karar ağacı sisteminin ticari versiyonu C5.0'dır [23,28]. J48 karar ağacı algoritması C4.5'in Weka'daki java uygulamasıdır (klon).

Tablo 1 – Nitelik Sayısı

Demografik veriler	Laboratuvar tetkikleri	Otonomik Testler	Cihaz ölçümleri		
			EKG	EKO	Holter
Cinsiyet	A1C	V1	S V1	SWth	dt sis
Yaş	AKS	V2	R V5	PWth	dt dias
Boy	Kreatinin	Vm		LVEDD	nt sis
Kilo	malb	DSD		LVESD	nt dias
BELC	LDL	AKH		ME	24h sis
BSA	ABI	OrH		MA	24h dias
BMI		HDKBF		MDT	mask HT
		PH		AoIVRT	%sisD
		SH		LVM	%diasD
		Ewing		LVMI	
				EF	
				LA	
				Ao	
7	6	10	2	13	9
2	2	6	1	10	0

J48 algoritması Enformasyon Kazancı Teorisine (Information Gain Theory) dayanarak, verilerden ilgili özellikleri seçmek için otomatik işlem yeteneğine sahiptir. Enformasyon kazancının en iyi olduğu noktadan örnekleri bölen yinelemeli algoritmadır. IF-THEN kurallarına dayalı bir karar ağacı ve “üyelik fonksiyon kümeleri -membership function sets” çıktısı verir [32]. Ağaç yapısı, denekleri bölme ve ağacın en iyi kök değişkeninin seçilmesi süreci ile başlayıp yukarıdan aşağıya doğru inşası gerçekleştirilmektedir [23]. J48, anlamlı olmayan diğer bir deyişle zayıf dalları kesmek için etkin bir budama işlemi yapabilmektedir. Bunun nedenlerinden biri, karar ağaçlarının amacının veri keşfetmek değil, veriler üzerinde basit bir sınıflandırma modeli oluşturmak olmasındandır. Başka bir nedeni ise, karar ağacının sezgisel

(heuristic) araması çok daha kaliteli kuralları bulmasını engellemektedir. Karar ağaçları, ağaç yapımında sadece bir yol takip ederler ve dolayısıyla alternatif yollar boyunca olabilecek daha iyi kuralları kaçırabilmektedirler. Son zamanlarda, bir varyant karar ağacı algoritması, yüksek-kazanç-bölme ağaç yöntemi, iş zekâsı yüksek ve faydalı kalıpları keşfetmek için önerilmiştir [33]. Tıbbi veriler üzerinde uygulanması araştırılmalıdır.

Sınıflandırıcı özneteliği seçmek için öncelikle sistemin entropisi hesaplanmıştır[34]. Sistemin entropisi S;

$$\text{Dipper}(N_{pos}) = 32$$

$$\text{Non_Dipper}(N_{neg}) = 24$$

$$\text{Toplam Veri Sayısı (Toplam)} = N_{pos} + N_{neg} = 56 \quad (1)$$

$$p_+ = N_{pos}/\text{Total} \quad (2)$$

$$p_+ = 32/56 = 0.571$$

$$p_- = N_{neg}/\text{Total} \quad (3)$$

$$p_- = 24/56 = 0.4285$$

$$H(S) = -p_+ \log_2(p_+) - p_- \log_2(p_-) = 0.9853 \quad (4)$$

Eğitim için ayrılan deneklere ait enformasyon kazancı kullanılarak, en iyi sınıflandırıcı tek bir öznetelik hesaplandı.

Herbir nitelik değeri için;

$$Gain(S, A_i) = H(S) - \sum_{v \in \text{Values}(A_i)} P(A_i = v) H(S_v) \quad (5)$$

En iyi enformasyon kazancı olan nitelik kök olarak belirlenmektedir.

3. Bulgular

Bu deneyde 56 deneğin % 66'sı eğitim için ayrılmıştır. Böylece, J48 algoritması 19 denek üzerinde test edilmiştir. Şekil 1'de ortaya çıkan karar ağacı modeli görülmektedir. Tablo 2'de test edilen her bir deneğe ait gerçek class değeri ile programın tahmini olan class değeri belirtilerek, hatalı sonuçlar işaretlenmiştir.

yok anlamına gelmektedir. Değer = 0 ise, ilişkinin tamamen tesadüfen olduğuna işaret etmektedir. κ istatistiği subjektif değerlendirmedir. 0.4751 değeri ise, Silman Tablosuna göre ortalama başarı anlamına gelmektedir [35,36].

$$\kappa = \frac{(a+d) - [(a+c)(a+b) + (b+d)(c+d)]/n}{n - [(a+c)(a+b) + (b+d)(c+d)]/n} \quad (6)$$

$$\text{Ortalama Mutlak Hata} = \frac{|p_1 - a_1| + \dots + |p_n - a_n|}{n} \quad (7)$$

$$\text{Ortalama Karesel Hatanın karekökü} = \sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}} \quad (8)$$

$$\text{Bağıl mutlak hata} = \frac{|p_1 - a_1| + \dots + |p_n - a_n|}{|a_1 - \bar{a}| + \dots + |a_n - \bar{a}|} \quad (9)$$

$$\text{Root relative squared error} = \sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{(a_1 - \bar{a})^2 + \dots + (a_n - \bar{a})^2}} \quad (10)$$

Tablo 3 – Sonuç Değerlendirmesi

	TP Oranı	FP Oranı	ROC Alanı	Class
	0.7	0.222	0.739	Non-Dipper
	0.778	0.3	0.739	Dipper
Ağırlıklı Ort.	0.737	0.259	0.739	

Tablo 4 –Hata Matrisi

Hata Matrisi		
a	b	sınıflandırma
7	3	a = Non-Dipper
2	7	b = Dipper
		Doğruluk % = (7+7)/(7+2+3+7) = 73.6842

Aşağıda belirtilen kriterler üzerinden hata matrisindeki hata sayılarına bakılarak, karar ağacı modeli performansı değerlendirilir. Kriterler:

True Positive (TP): Tıp uzmanının Holter bulgularına göre non-dipper teşhisi koyduğu denek, program tarafından da hasta (nondipper) tesbit edildi.

True Negative (TN): Tıp uzmanının Holter bulgularına göre dipper teşhisi koyduğu denek, program tarafından da normal (dipper) tesbit edildi.

False Positive (FP): Tıp uzmanının Holter bulgularına göre dipper teşhisi koyduğu denek, program tarafından hasta (nondipper) tesbit edildi.

False Negative (FN): Tıp uzmanının Holter bulgularına göre non-dipper teşhisi koyduğu denek, program tarafından normal (dipper) tesbit edildi.

Sağlam kişilerden oluşan denek grubu verildiğinde testin pozitif olma olasılığı, testin duyarlılığıdır. Duyarlılık formülü:

$$\begin{aligned} \text{Duyarlılık} &= (\text{TP} / (\text{TP} + \text{FN})) & (20) \\ &= (7 / (7 + 3)) \\ &= 0.70 \\ &= \%70 \end{aligned}$$

Bir testin Belirlilik oranı, hastalığı olan deneklerde testin negative olma olasılığıdır. Belirlilik için formül:

$$\begin{aligned} \text{Belirlilik} &= (\text{TN} / (\text{TN} + \text{FP})) & (21) \\ &= (7 / (7 + 2)) \\ &= 0.7777 \\ &= \%78 \end{aligned}$$

4. Tartışma

Bu çalışma geliştirildiğinde kardiyovasküler açıdan daha yüksek risk altındaki kişiler pratik ve hızlı bir yöntemle anlaşılabilir, gereken tedavileri daha kolaylıkla planlanabilecektir. Böylece; hasta yaşam kalitesi ve maliyetler açısından var olan olumsuzlukları ortadan kaldırmaya yönelik fayda sağlanabilecektir.

Veri madenciliği yönteminde denek sayısının dolayısıyla tıbbi veri sayısının çokluğu çalışmanın güvenilirliği açısından önem arz etmektedir. Bu bağlamda, bu pilot çalışmanın, sürdürülen disiplinler arası projenin ilk bölümünü oluşturduğunu belirtmek önemlidir.

5. Sonuç

Literatürde “Non-dipping” özelliğın öngörülmesine, bu yöntemle sınıflandırılmasına yönelik benzer bir çalışmaya rastlanılmamıştır. Disiplinler arası yürütülen bu çalışma özgün bir çalışma olup, sonuçları itibarıyla tıbbi olarak nondipper/dipper sınıflaması için holter cihazı verileri gerekmebileceğini, hasta yaşam kalitesi ve maliyetler açısından ortaya çıkan olumsuzluklara çözüm sunulabileceğinin ilk sinyalini vermiştir. Çalışmanın denek sayısı artırılarak devam ettirilmesi ve yapay sinir ağları algoritmaları ile doğruluk, duyarlılık ve belirlilik kriterlerinin %80’in üzerine çıkartılabilecek şekilde yeni modeller oluşturulması gerekliliğinin sonucuna varılmıştır.

6. Teşekkür/not

Bu çalışmaya olan katkılarından dolayı Prof. Dr. Hakan Tezcan, Doç. Dr. Hayriye Korkmaz ve Dr. Davut Bilgin'e teşekkür ederiz.

Bu çalışma; Marmara Üniversitesi Bilimsel Araştırma Projeleri Komisyonu tarafından FEN-C-DRP-110908-0225 proje numarasıyla desteklenmiştir.

Tip 2 diyabet ve normotansif hasta profilinden oluşan 65 kişilik bu denek grubu için 08.05.2009 tarih/MAR-YÇ-2009-0166 numaralı Marmara Üniversitesi Tıp Fakültesi Klinik Araştırma Etik Kurulu izni alınmıştır.

7. Kaynakça

- [1] Dr. Türker H.: "Hipertansif ve Normotansif Tip 2 Diyabet Hastalarında İnsülin Direncinin Karşılaştırılması", Uzmanlık Tezi, İstanbul- 2006
- [2] Wild S.; Roglic G.; Green A.; Sicree R.; King H.: "Global Prevalence of Diabetes: Estimates for the year 2000 and projection for 2030", *Diabetes Care* 2004, 27: 1047-1053.
- [3] Weed L. *Medical Records, Medical Education and Patient Care*. 2nd ed. Cleveland: Case Western Reserve University Press, 1971.
- [4] Satman İ.: "Diabetes Mellitus'un Epidemiyolojisi", Yenigün M., Altuntaş Y. (eds), Her Yönüyle Diabetes Mellitus, 2. Baskı, Nobel Tıp Yayınevi İstanbul 2001:69-84, ISBN: 9754200521
- [5] King H.; Rewers M.: "Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults: the WHO AdHoc Diabetes Reporting Group", *Diabetes Care* 1993, 16: 157-177.
- [6] Dr. Bozkurt B. O.: "Haseki Eğitim Ve Araştırma Hastanesi İç Hastalıkları Nefroloji Polikliniğinde Takip Edilen Aşık Nefropatili Tip 2 Diyabetik Hastalarda Renal Prognozun Göstergeleri", Uzmanlık Tezi, İstanbul 2007
- [7] Amos A.F.; McVarty D.J.; Zimmet P.: "The rising global burden of diabetes and its complications: Estimates and projections to the year 2010", *Diabetic Med* 1997, 14:7-55
- [8] Zimmet P.; Dowse G.; Finch C.; King H.: "The epidemiology and natural history of NIDDM- lessons from South Pasific.", *Diabetes Metabo. Rev.* 1990, 6: 91-124.
- [9] Dr. Morgül G.: "Tip 2 Diyabetes Mellitus Hastalarında Serum Lipoprotein Düzeyleri", Uzmanlık Tezi, Adana - 2008
- [10] Satman İ.: Epidemiology Study. Population-based study of diabetes and related risk characteristics of Turkey, Results of the Turkish Diabetes Care 2002, 25: 1551-1556.
- [11] Satman İ.; Şengül A.M.; Uygur S.; Salman S.; Baştar İ.; Sargin M.; Tütüncü Y.; Karşıdağ K.; Dinççağ N.; The TURDEP group, *Diabetologia*, 2000,11:1.
- [12] Satman İ.; Yılmaz T.; The Turdep Group: "Population-Based Study Of Diabetes and Risk Characteristics in Turkey Results of The Turkish Diabetes Epidemiology Study (Turdep)", *Diabetes Care*, September 2002, 25(9): 1551- 1556
- [13] Onat A.; Hergenç G.; Uyarel H.; Günay C.; Özhan H.: "Prevalence, Incidence, Predictors and Outcome of Type 2 Diabetes in Turkey", *Anadolu Kardiyoloji Dergisi*, Aralık 2006, 6(2): 314 -321.
- [14] Dr. Taşkapan H.: "Hipertansiyon Tanısı Nasıl Konulmalıdır", 11.Ulusal Hipertansiyon ve Böbrek Hastalıkları Kongresi, 13-17 Mayıs 2009 Antalya.
- [15] Tzamou V.; Vyssoulis G.; Karpanou E.; Kyvelou S.M.; Gialermios T.; Stefanadis C.: "Serum Phosphate in White Coat Hypertensive Patients: Focus On Dipping Status and Metabolic Syndrome", *Journal of Hypertension*, 28, e-Supplement A, June 2010
- [16] Sitar-Taut D.A.; Sitar-Taut A.V.: "Overview on How Data Mining Tools May Support Cardiovascular Prediction", *Journal of Applied Computer Science & Mathematics*, Suceava, 2010, 8 (4).
- [17] Witten H. I.; Frank E.; Hall A. M.: *Data mining: practical machine learning tools and techniques.*—3rd ed., Elsevier Inc., 2011, ISBN 978-0-12-374856-0
- [18] <http://www.cs.waikato.ac.nz/ml/weka/> Erişim Tarihi: 01.04. 2011.
- [19] Sitar-Taut V. A.; Zdrenghea D.; Pop D.; Sitar-Taut. A.D.: "Using machine learning algorithms in cardiovascular disease risk evaluation", *Journal of Applied Computer Science & Mathematics*, 20095(3):29- 32.
- [20] SITAR-TĂUT D.A.; SITAR-TĂUT A.V.: Overview on How Data Mining Tools May Support Cardiovascular Disease Prediction, *Journal of Applied Computer Science & Mathematics*, Suceava, 2010,8 (4)
- [21] Safavian R.; Landgrebe D.: "A survey of Decision Tree Classifier Methodology", *IEEE Transactions on Systems, Man and Cybernetics*, May/June 199121(3): 660-674.

- [22] Pooja A. P.; Jayanth J.; Dr. Shivaprakash K.: "Classification of RS data using Decision Tree Approach", International Journal of Computer Applications June 2011, 23(3): 0975 – 8887
- [23] Li J.; Fu A. W.; Fahey P.: "Efficient discovery of risk patterns in medical data." Artificial Intelligence in Medicine, 2009,45: 77–89
- [24] Kononenko I.: "Machine learning for medical diagnosis: history, state of the art and perspective.", Artificial Intelligence in Medicine 2001,1:89–109.
- [25] Zhou Z.; Jiang Y.: "Medical diagnosis with C4.5 rule preceded by artificial neural network ensemble.", IEEE Transactions on Information Technology in Biomedicine 2003,1:37–42.
- [26] Ordonez C.: "Comparing association rules and decision trees for disease prediction." In: Xiong L, Xia Y, editors. Proceedings of the international workshop on healthcare information and knowledge management, New York: ACM, 2006: 17–24.
- [27] Gaga L.; Moustakis V.; Vlachakis Y.; Charissis G.: "ID+: Enhancing Medical Knowledge Acquisition with Machine Learning", Applied Artificial Intelligence, Taylor & Francis, 1996,10:79- 94.
- [28] Quinlan J.R.: Induction of Decision Trees, Machine Learning, 1986: 81-106
- [29] Hewett R.; Leuchner J.; Mooney S. D.; Klein T.E.: "Analysis of mutations in the colia1 gene with second order rule induction", International Journal of Pattern Recognition and Artificial Intelligence, World Scientific Publishing Company, 2003, 17(5):721-740
- [30] Quinlan J.R.: C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, San Mateo, CA,1993
- [31] Mazid M. M.; Shawkat A. B. M. A.; Tickle K. S.: Improved C4.5 Algorithm For Rule Based Classification, Recent Advances In Artificial Intelligence, Knowledge Engineering And Data Bases, ISBN: 978-960-474-154-0 Artificial Intelligence Series, Wseas Press, 2010.
- [32] Mollazade K.; Ahmadi H.; Omid M.; Alimardani R.; "An Intelligent Combined Method Based on Power", International Journal of Intelligent Technology, Spectral Density, Decision Trees and Fuzzy Logic for Hydraulic Pumps Fault Diagnosis, 2008, 3(4): 251-263.
- [33] Hu J.; Mojsilovic A.: "High-utility pattern mining: a method for discovery of high-utility item sets.", Pattern Recognition, 2007,11:3317–24.
- [34] Doğan Y.: "A Data Mining Based Target Classification for Tactical Underwater Sensor Networks", Master Thesis, Turkish Naval Academy, Naval Sciences and Engineering Institute, Department of Computer Engineering, 2004, Istanbul
- [35] Brennan P.; Şilman A.: "Statistical methods for assessing observer variability in clinical measures.", BMJ., Jun 1992, 6,304(6840):1491-4.
- [36] Demir Y. K.; Gerçek H.: "Flexible methods of calculating the transport type selection", itüdergisi/d mühendislik, December 2006 5(6): 61-73.

7. Sorumlu Yazarın Adresi

Dr. Aysun ALTIKARDEŞ, aaltikardes@marmara.edu.tr, 0216 418 25 06-609, 0505 739 80 00, Marmara Üniversitesi Teknik Bilimler Meslek Yüksekokulu Göztepe Kampüsü 34722 Kadıköy- İstanbul