

# Hipertansiyonun tahmini için çoklu tahmin modellerinin karşılaştırılması (Sinir ağları, lojistik regresyon ve esnek ayırma analizleri)

*Comparison of multiple prediction models for hypertension  
(Neural network, logistic regression and flexible discriminant analyses)*

Dr. Mevlüt Türe, Dr. İmran Kurt, Dr. Ebru Yavuz, Dr. Turhan Kürüm”

Trakya Üniversitesi Tıp Fakültesi Biyoistatistik ve “Kardiyoloji Anabilim Dalları, Edirne

## ÖZET

**Amaç:** Bu çalışmada, primer hipertansiyonun tahminlenmesinde lojistik regresyon analizi (LR, Logistic Regression Analysis), esnek diskriminant analizi (EAA, Flexible Discriminant Analysis) ve sinir ağları (SA, Neural Networks) yöntemlerinin karşılaştırılması amaçlandı.

**Yöntemler:** Aile hikayesi, lipoprotein A, trigliserid, sigara kullanımı ve vücut kitle indeksi tahminleyici değişkenler olarak ele alındı. Veriler, 2001 yılında Trakya Üniversitesi Tıp Fakültesi Kardiyoloji Kliniğinden elde edildi. Kontrol ve hipertansiyonlu bireylerden oluşan grupların tahminlenmesinde LR, EAA ve SA yöntemleri kullanıldı. Modellerin performansları ROC analiz ile karşılaştırıldı.

**Bulgular:** Bütün modellerin ROC eğrisi altındaki alanları 0.793-0.984 aralığında yer aldı. Sinir ağlarının duyarlılık, özgüllük ve doğruluk oranları %90'dan daha yüksek bulundu. Sinir ağları ve LR ile SA ve EAA'nın ROC eğrisi altında kalan alanları istatistiksel olarak farklı bulundu (sırasıyla,  $p<0.001$  ve  $p<0.001$ ). Buna karşın EAA ve LR'nin ROC eğrisi altında kalan alanları istatistiksel olarak farklı bulunmadı ( $p>0.05$ ).

**Sonuç:** Aile hikayesi, lipoprotein A, trigliserid, sigara kullanımı ve vücut kitle indeksi değişkenlerinin, kontrol ve hipertansiyonlu hasta gruplarını tahmin etmede kullanılabileceğine ve SA'nın performansının LR ile EAA'dan istatistiksel olarak daha iyi olduğuna karar verildi. (*Anadolu Kardiyol Derg 2005; 5: 24-8*)

**Anahtar kelimeler:** Lojistik regresyon analizi, sinir ağları, esnek ayırma analizi, ROC eğrisi

## ABSTRACT

**Objective:** In this study, we compared performances of logistic regression analysis (LR), flexible discriminant analysis (EAA) and neural networks (SA) in prediction of primary hypertension.

**Methods:** Predictor variables were family history, lipoprotein A, triglyceride, smoking and body mass index. The data were collected from Cardiology Clinic of Trakya University Medical Faculty in Turkey, 2001. Logistic regression analysis, flexible discriminant analysis and neural networks were used for prediction of control and hypertension groups. Comparison of the performance of all models was done using receiver operating characteristic (ROC) curve analysis.

**Results:** All models had areas under the ROC curve in the range of 0.793-0.984 and SA had sensitivity, specificity, and accuracy greater than 90% at ideal threshold. ROC curve areas of SA and LR, and SA and EAA were statistically different ( $p<0.001$  and  $p<0.001$  respectively), while ROC curve areas of EAA and LR did not differ ( $p>0.05$ ).

**Conclusion:** We concluded that family history, lipoprotein A, triglyceride, smoking and body mass index variables can be used for prediction of control and hypertension groups with statistically better performance of SA over LR and EAA. (*Anadolu Kardiyol Derg 2005; 5: 24-8*)

**Key words:** Logistic regression analysis, neural networks, flexible discriminant analysis, ROC curve

## Giriş

Hipertansiyon, ülkemizdeki önemli sağlık sorunlarından biridir. Hipertansiyon, normal düzeyin üzerine çıkan kan basıncı seviyesi olarak tanımlanır. Hipertansiyon erken tanı ile kontrol altına alınabilen; geç kalındığında kalp, beyin, böbrekler ve gözlerde hasara

neden olabilen ciddi bir hastalıktır (1-4). Genetik faktörlerin kan basıncı yüksekliğinde önemli bir rol oynadığı düşünülmektedir (5-11). Hipertansiyon, poligenlerle kuşaklar arası geçiş göstermesi yanında çevresel faktörlerle etkileşim sonucunda kan basıncının yükselmesiyle ortaya çıkmaktadır. Hipertansiyonun gelişimi dolayısıyla kan basıncının yükselmesinde, genetik yatkınlığın yanı sıra çev-

resel faktörlerden kaynaklanan birçok risk faktörü vardır (1-11). Çalışmamızda risk faktörü olarak aile hikayesi, lipoprotein A, trigliserid, sigara kullanımı ve vücut kitle indeksi değişkenleri yer almıştır (1-11).

Bu çalışmada, kontrol ve primer hipertansiyonlu bireylerden oluşan grupların tahminlenmesinde lojistik regresyon analizi (LR), esnek ayırma analizi (EAA) ve sinir ağları (SA) yöntemlerinin karşılaştırılması amaçlandı.

## Yöntemler

**Lojistik regresyon analizi:** Lojistik regresyon analizi, bağımlı değişkenin kategorik ve bağımsız değişkenlerin karışık ölçekli olması durumunda belirli bir dağılım varsayımına bağlı kalmaksızın bağımlı değişken ile bağımsız değişkenler arasındaki neden-sonuç ilişkisinin belirlenmesinde ve aynı zamanda bağımsız değişkenlerin etkilerine dayanarak verilerin sınıflandırılmasında kullanılan bir yöntemdir (12-14).

Bağımsız değişken olan "p" için LR modeli

$$P(y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}}$$

olup  $\beta_0, \beta_1, \dots, \beta_p$ , regresyon katsayılarıdır (13,14).

**Esnek ayırma analizi:** Doğrusal ayırma analizi (DAA), grupları ortalamalarına göre ortak ortalamadan farklı olmalarını sağlayacak bir ayırma kriteri geliştirmeyi amaçlayan bir yöntemdir. Bu nedenle veri setlerine DAA uygulanabilmesi için; veri matrisinin çok değişkenli normal dağılım göstermesi, değişkenlerin varyans ve kovaryanslarının homojen olması, değişkenlerin ortalamaları ve varyansları arasında ilişki bulunmaması, değişkenler arasında çoklu bağımlılık bulunmaması, veri matrisinin grupların birbirlerinden ayrılmasında rol oynamayacak gereksiz değişkenler içermemesi varsayımlarının sağlanması gerekir (12,15).

Esnek ayırma analizi; DAA'nın varsayımlarının getirmiş olduğu kısıtlamalar nedeniyle ortaya çıkan problemleri ortadan kaldırarak birimleri sınıflandırmada kullanılan, DAA'nın genelleştirilmiş parametrik olmayan şeklidir (16-20).

**Sinir ağları:** Sinir ağları, elle çözüm imkanı vermeyen son derece karmaşık yapıya sahip, insan beynindeki sinir ağları gibi çalışarak en zor problemlerin çözümünü sağlayan ve değişken yapısı konusunda herhangi bir kısıtlama getirmeksizin değişkenler arasındaki ilişkiyi ortaya koyan çok esnek bir yöntemdir. Sinir ağları, nöronlar olarak bahsedilen basit hesap hücrelerinin birbirleriyle bağlantılarını kullanarak insan beyinde olduğu gibi bilgiyi kaybetmeden hedefe doğru minimum hata ile ulaşılmasını sağlar (21-27).

Sinir ağları, ağ hatasının optimal değerini bulmak için geriye doğru yayılma algoritmasını kullanarak aşama aşama en iyi tahmine ulaşmaya çalışır (22-25). Basit bir ağda, p girdi vektörü ile W ağırlık matrisi çarpılır ve bu çarpıma b sapmasının eklenmesiyle f dönüşüm fonksiyonunun girdisi elde edilir. Ağ, elde edilen bu yeni girdiyi f dönüşüm fonksiyonu yardımıyla ağ içinden iletir a çıktısını oluşturur (22-24,26).

Sinir ağlarında kullanılan hiperbolik tanjant sigmoid dönüşüm fonksiyonu

$$f(\sum wp+b) = \frac{2}{1 + e^{-2(\sum wp+b)}} - 1 \text{ biçimindedir (21,22).}$$

**ROC (receiver operator curve) eğrisi yöntemi:** ROC eğrisi, tanı testlerinin performanslarının değerlendirilmesi ve kıyaslanması için yaygın olarak kullanılan bir yöntemdir. ROC eğrisi, duyarlılık ve özgüllük oranlarını kullanarak birimleri sınıflarına ayıran en uygun kesim noktasını belirler. Sınıflamanın doğruluğu, ROC eğrisi altında kalan alanın büyüklüğüne bağlıdır. ROC eğrisi altında kalan alan, doğru tanı testinin seçiminde kullanılan çok popüler bir ölçüdür. Tanı testinin, birimleri gruplara ayırmadaki performansının belirlenmesinde ve başka tanı testleriyle karşılaştırılmasında ROC eğrisi yönteminden sıklıkla yararlanılmaktadır (28-30).

Hanley ve McNeil 1983 yılında ROC eğrisi altında kalan alanlardan yararlanarak iki tanı testinin performansını karşılaştıran bir yöntem geliştirmişlerdir. Bu yöntemde, aynı birimler üzerinde farklı iki tanı testi uygulanması durumunda ROC eğrisi altında kalan alanlar karşılaştırılır (30).

## Veriler

Geriye dönük ve olgu-kontrol araştırması olarak gerçekleştirilen çalışmamız, 2001 yılında Trakya Üniversitesi Tıp Fakültesi Kardiyoloji Polikliniğine gelen 65 yaşın altında 236 hasta sistolik kan basıncı (SKB) 140 mmHg ve diyastolik kan basıncı (DKB) 90 mmHg olan primer hipertansiyonlu) ve 123 bireylik kontrol gruplarından oluşmaktadır. Hipertansif hasta grubu, yeni tanı alan ancak tedavi almayan ve daha önce tanı konulan ancak tedavi almayan ya da tedaviye ara vermiş olan, farklı zamanlarda en az üç kez ölçülen SKB 140 mmHg ve/veya DKB 90 mmHg olan 236 hastadan oluşturuldu. Kontrol grubu ise SKB 100-135 mmHg arasında değişen, ortalama ve standart sapması 118.75 8.29 mmHg ve DKB 60-88 mmHg arasında değişen, ortalama ve standart sapması 78.37 5.11 mmHg olan toplam 123 sağlıklı bireyden oluşturuldu.

İstatistiksel analizlerde LR için Minitab Release 13.32, EAA için R 1.81, SA için NeuroSolutions 4.0 ve ROC eğrisi altında kalan alanların karşılaştırılmasında MedCalc paket programları kullanılmıştır.

## Bulgular

Hasta grubunun %28.2'si kadın, %71.8'si erkek ve kontrol grubunun %36.8'i kadın %63.2'si erkek olmakla birlikte Ki-kare analizi sonucunda hasta ve kontrol grubu arasında cinsiyet bakımından anlamlı fark olmadığı belirlendi (p>0.05). Yaş değişkeninin normal dağılıma uygunluğu Kolmogorov-Smirnov testi ile incelendi. Normal dağılıma uygun olduğu belirlenen yaş değişkenine bağımsız gruplarda "t" testi uygulanması sonucunda gruplar arasında anlamlı fark olmadığı (p>0.05), hasta grubunun yaş ortalamasının 48.21±8.63 yıl ve kontrol grubunun yaş ortalamasının 46.51±8.22 yıl olduğu bulundu. Aile hikayesi, lipoprotein A, trigliserid, sigara kullanımı ve vücut kitle indeksi bağımsız değişkenler olarak ele alındı. Bu değişkenlere ilişkin tanımlayıcı istatistikler Tablo 1'de verilmiştir.

Bu çalışma, 2 aşamadan oluşmaktadır. Birinci aşama SA'nın çalışma ve test sonuçlarının incelenmesi; 2. aşama ise SA ile LR ve EAA'nın karşılaştırılmasıdır.

**Sinir ağlarının çalışma ve test sonuçlarının karşılaştırılması:** Sinir ağlarında parametre tahmini için kullanılan veri setinden farklı bir set üzerinde modelin geçerliliğini test etmek için çalışma seti dışında bir test seti seçilir. Genellikle toplam birim sayısının en az %10'u test seti olarak alınır (31). Çalışmamızda, test seti sonuç-

larının yüzdelik oranlarla gösterilebilmesi için toplam birim sayısının %23'ü test setine alındı. Böylece SA modeli oluşturulmadan önce veri seti, 359 birimin %77'si olan 276 birimlik (181 hasta ve 95 kontrol) çalışma ve 359 birimin %23'ü olan 83 birimlik (55 hasta ve 28 kontrol) test seti olmak üzere rastlantısal olarak iki gruba ayrıldı. Sinir ağları, 15 gizli nöronlu (hiperbolik tanjant sigmoid dönüşüm fonksiyonlu) ve 1 ikili (hasta-kontrol) çıktılı (hiperbolik tanjant sigmoid dönüşüm fonksiyonlu) ağdan oluşmaktadır. Sinir ağlarında hatayı minimize etmek amacıyla geriye doğru yayılma algoritması kullanıldı. Sinir ağları modeli ile elde edilen sınıflandırma sonuçları Tablo 2'de verilmiştir. Sinir Ağları modelinin hastaların %97.1'ini, kontrollerin ise %92.7'sini doğru sınıflandırdığı bulundu.

Çalışma setinden elde edilen SA modeli test setine uygulandı. Çalışma ve test setinin performansları karşılaştırıldığında sonuçlarının benzer oldukları görüldü (Şekil 1, Tablo 3).

**Sinir ağlarının LR ve EAA ile karşılaştırılması:** Lojistik regresyon analizi ve EAA sonucunda elde edilen parametre tahminleri ve LR ile parametrelerin anlamlılık testi sonuçları ve "odds" oranı (OR) değerleri Tablo 4'de verilmiştir. Lojistik regresyon analizi sonuçlarına göre lipoprotein A arttıkça hipertansiyon hastası olma oranının 1.014 kat arttığı, trigliserid arttıkça hipertansiyon hastası olma oranının 1.008 kat arttığı, vücut kitle indeksi arttıkça hipertansiyon hastası olma oranının 1.092 kat arttığı, aile hikayesinde primer hipertansiyon olanların olmayanlara göre hipertansiyon hastası olma oranının 6.815 kat daha fazla olduğu ve sigara kullananların kullanmayanlara göre hipertansiyon hastası olma oranının 2.141 kat daha fazla olduğu bulundu. Lojistik regresyon analizi için model uyumu Hosmer-Lemeshow testi ile incelendiğinde, LR modelinin hasta ve kontrol gruplarını ayırmada yeterli bir model olduğu bulundu ( $\chi^2=3.031$ ,  $sd=8$  ve  $p>0.05$ ).

Lojistik regresyon analizi modeli ile elde edilen sınıflandırma sonuçları Tablo 5'de verilmiştir. Lojistik regresyon analizi modelinin hastaların %85.4'ünü, kontrollerin ise %60.2'sini doğru sınıflandırdığı bulundu. Esnek ayırma analizi modeli ile elde edilen sınıflandırma sonuçları ise Tablo 6'da verilmiştir. Esnek ayırma analizi mo-

**Tablo 1. Bağımsız değişkenlerin gruplara göre tanımlayıcı istatistikleri**

Bağımsız Değişkenler	Hasta Grubu n=236	Kontrol Grubu n=123
Aile hikayesinde hipertansiyon (%)	77.7	33.6
- Olan	22.3	66.4
- Olmayan		
Lipoprotein A (mg/dl)	33.42±24.14	26.60±16.88
Trigliserid (mg/dl)	164.13±82.22	130.07±59.18
Sigara kullanımı (%)		
- İçen	23.8	32.7
- İçmeyen	76.2	67.3
Vücut kitle indeksi (kg/m <sup>2</sup> )	28.80±3.62	27.28±3.64

**Tablo 2. Sinir ağları ile elde edilen çalışma seti sınıflandırma tablosu**

Tahmin Sonucu	Gerçek Durum			Toplam
	Hasta	Kontrol		
<b>Hasta</b>	176 (%97.1)	7 (%7.3)		183
<b>Kontrol</b>	5 (%2.9)	88 (%92.7)		93
<b>Toplam</b>	181	95		276

delinin hastaların %82'sini, kontrollerin ise %63.7'sini doğru sınıflandırdığı bulundu.

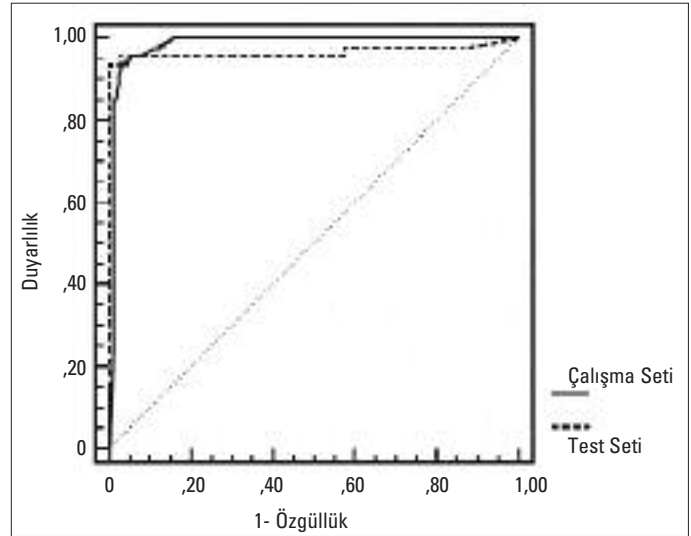
Lojistik regresyon analizi, EAA ve SA'nın ROC eğrisi altındaki alanları 0.793-0.979 aralığında elde edildi (Tablo 7). Lojistik regresyon analizi, EAA ve SA'nın ROC eğrisi altında kalan alanları Şekil 2'de gösterilmiştir. Sinir Ağları ile LR ve SA ile EAA'nın ROC eğrisi altında kalan alanları istatistiksel olarak farklı bulundu (sırasıyla,  $p<0.001$  ve  $p<0.001$ ). Esnek ayırma analizi ile LR'nin ROC eğrisi altında kalan alanları istatistiksel olarak farklı bulunmadı ( $p>0.05$ ). Ayrıca eğri altında kalan alanların standart hataları incelendiğinde, SA'nın standart hatasının diğerlerine göre daha küçük olduğu görüldü.

## Tartışma

Özellikle yaş ilerledikçe morbidite ve mortalite nedenleri olan kalp yetersizliği, inme, böbrek yetersizliği, koroner kalp hastalığı ve periferik damar hastalığına yol açan hipertansiyonu etkileyen risk faktörlerini belirleyici ve hipertansiyon hastalarını sağlam bireylerden ayırıcı istatistiksel modellerin geliştirilmesine yönelik çalışmalar önem kazanmaktadır.

Hipertansif hastalarda, hipertansiyon dışındaki kardiyovasküler risk faktörlerine de rastlanmaktadır. Çalışmamızda hipertansiyonun tahmininde kardiyovasküler risk faktörleri arasında yer alan lipoprotein A, trigliserid, obezite, aile öyküsü, sigara kullanımı faktörleri kullanılmıştır.

Sigara kullanımı, kan basıncını yükselten önemli bir risk faktörü-



**Şekil 1. Çalışma ve test setlerinin ROC eğrileri**

**Tablo 3. Çalışma ve test setlerinin duyarlılık ve özgüllük oranları, ROC eğrisi altında kalan alanları, standart hataları ve ROC eğrisi altında kalan alanın güven aralığı**

	Çalışma Seti	Test Seti
Duyarlılık (%)	97.1	96.6
Özgüllük (%)	92.7	93.8
ROC eğrisi altında kalan alan	0.984	0.967
Standart hata	0.006	0.019
ROC eğrisi altında kalan alanın %95 güven aralığı	0.964 - 0.995	0.901 - 0.994

dür (32). Hipertansiyon dışında koroner arter hastalığı, inme, beyin kanaması, kanser, ani ölüm ve akciğer hastalığı riskini artırır. Çalışmamızda da sigara kullanımının hipertansiyonun tahmininde önemli bir risk faktörü olduğu ve sigara kullananların kullanmayanlara göre hipertansiyon hastası olma oranının daha yüksek olduğu bulundu.

Primer hipertansiyon, paleolitik genlerle taşınan ve yaşam koşulları ile görülme sıklığı giderek artan önemli bir halk sağlığı sorunudur (33,34). Anne, baba veya yakın kanbağı olan akrabalarda hipertansiyon olması, hipertansiyon oluşma riskini artırmaktadır. Aile fertleri arasında erken yaşta kalp krizi veya felç geçirenler bulunuyorsa, diğer aile üyelerinin de kan basınçlarını sık sık kontrol ettirmeleri erken tanı için önemlidir. Kalıtsal özelliklerin hipertansiyona katkısı, %30-60 gibi önemli bir orandır (35). Hipertansiyon gelişiminde genetik yatkınlığın önemli rolü çalışmamızda da aile hikayesinde hipertansiyon olanların olmayanlara göre hipertansiyon hastası olma oranlarının daha fazla olduğu şeklinde ortaya çıkmıştır.

Yapılan çalışmalarda hipertansif hastalarda trigliserid düzeylerinin yüksek olduğu bildirilmiştir (36,37). Çalışmamızda da hipertansiyonun tahmininde trigliserid düzeyinin önemli etkisinin olduğu ve trigliserid düzeyi arttıkça hipertansiyonlu hasta olma oranının da arttığı bulundu.

Obezite, kan basıncı yükselmesinde etkili olan risk faktörlerinden biridir (38). Yapılan taramalarda ideal ağırlıklarının %20'sinden fazla şişman olan kimselerde, hipertansiyon gelişme riskinin yüksek olduğu anlaşılmıştır (35). Çalışmamızda da vücut kitle indeksinin hipertansiyon hastalığında önemli bir risk faktörü olduğu ve vücut kitle indeksi arttığında hipertansiyon hastası olma oranının da arttığı bulundu.

Lipoprotein A, kardiyovasküler risk faktörlerinden biridir (37,39,40). Bu nedenle, hipertansiyon ile lipoprotein A arasındaki ilişki, hipertansif hastaların kardiyovasküler hastalığa yatkınlığı ve lipoprotein A'nın kardiyovasküler hastalık oluşumuna katkısı nedeniyle ortaya çıkmaktadır. Çalışmamızda hipertansiyonun tahmininde lipoprotein A düzeyinin önemli etkisinin olduğu ve lipoprotein A düzeyi arttıkça hipertansiyon hastası olma oranının da arttığı bulundu.

Çalışmamızda lipoprotein A, trigliserid, sigara kullanımı, aile hikayesi, vücut kitle indeksi değişkenlerinin hipertansiyon üzerinde önemli etkilerinin olduğu ve bu risk faktörlerinin birimleri hasta ve kontrol gruplarına ayırmada kullanılacak modelleme yöntemlerinde birlikte ele alınmaları gerektiğine karar verildi.

Hipertansif hasta ve kontrol gruplarının ayrılmasında LR, EAA ve SA'nın performanslarının karşılaştırıldığı çalışmamızda, bütün modellerin ROC eğrisi altındaki alanları 0.793-0.979 aralığında elde edildi. Modeller ROC eğrisi altında kalan alanlar bakımından karşılaştırıldığında SA'nın ROC eğrisi altında kalan alanı, LR ve EAA'nın alanlarından istatistiksel olarak daha büyük bulundu. Sınır ağlarında duyarlılık, özgüllük ve doğruluk oranları %90'dan daha büyük bulundu.

Sonuç olarak, aile hikayesi, lipoprotein A, trigliserid, sigara kullanımı ve vücut kitle indeksi değişkenlerinin, kontrol ve hipertansiyonlu hasta gruplarını tahmin etmede kullanılabileceğini, SA'nın grupları sınıflandırma gücünün klasik çok değişkenli tekniklere göre daha iyi olduğunu ve sınıflandırma problemlerinde SA tercih edilerek modelleme çalışmalarının yapılması gerektiğini söyleyebiliriz.

**Tablo 4. Lojistik regresyon ve esnek ayırma analizi sonucunda elde edilen bağımsız değişkenlere ilişkin parametre tahminleri**

Bağımsız Değişkenler	EAA		Standart Hata	Wald	LR		OR	OR'nin % 95 Güven Aralığı	
	$\beta_i$	$\beta_i$			sd	p		Alt Sınır	Üst Sınır
Lipoprotein A (mg/dl)	0.011	0.014	0.007	4.133	1	p<0.05	1.014	1.000	1.027
Trigliserid (mg/dl)	0.005	0.008	0.002	12.194	1	p<0.001	1.008	1.003	1.012
Vücut Kitle İndeksi (kg/m <sup>2</sup> )	0.075	0.088	0.038	5.325	1	P<0.05	1.092	1.013	1.177
Aile Hikayesi	1.853	1.919	0.279	47.476	1	p<0.001	6.815	3.948	11.764
Sigara Kullanımı	0.640	0.761	0.312	5.970	1	p<0.05	2.141	1.163	3.943
Sabit	-5.545	-5.028	1.175	18.307	1	p<0.001	0.007		

EAA: esnek ayırma analizi, LR: lojistik regresyon analizi, OR: "odds" oranı

**Tablo 5. Lojistik regresyon analizi ile elde edilen sınıflandırma tablosu**

Tahmin Sonucu	Gerçek Durum			Toplam
	Hasta	Kontrol		
Hasta	155 (%85.4)	38 (%39.8)		193
Kontrol	26 (%14.6)	57 (%60.2)		83
Toplam	181	95		276

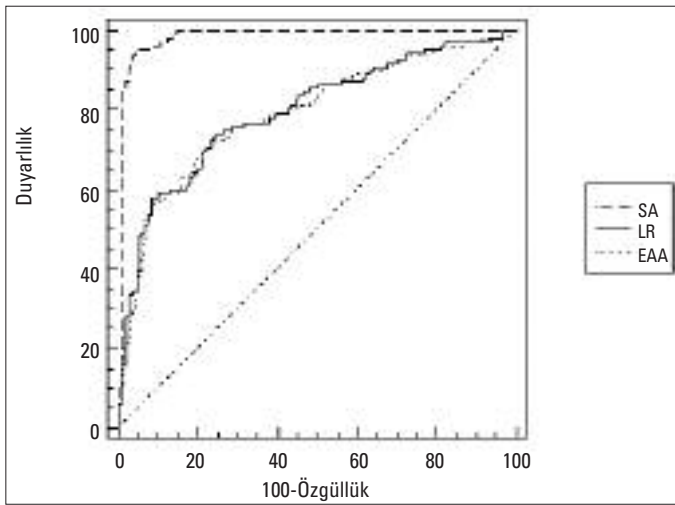
**Tablo 6. Esnek ayırma analizi ile elde edilen sınıflandırma tablosu**

Tahmin Sonucu	Gerçek Durum			Toplam
	Hasta	Kontrol		
Hasta	148 (%82.0)	34 (%36.3)		182
Kontrol	33 (%18.0)	61 (%63.7)		94
Toplam	181	95		276

**Tablo 7. Duyarlılık, özgüllük, pozitif tanımlama, negatif tanımlama ve doğruluk oranları, ROC eğrisi altında kalan alanlar, kesim noktaları ve standart hatalar.**

	LR	EAA	SA
Duyarlılık (%)	85.4	82.0	97.1
Özgüllük (%)	60.2	63.7	92.7
Pozitif tanımlama oranı (%)	79.6	80.5	98.5
Negatif tanımlama oranı (%)	69.4	66.1	88.0
Doğruluk (%)	76.5	75.5	94.4
ROC eğrisi altında kalan alan	0.800	0.793	0.984
Kesim noktası	0.499	0.498	0.433
Standart hata	0.028	0.027	0.006

EAA: esnek ayırma analizi, LR: lojistik regresyon analizi, SA: sınır ağları



**Şekil 2. LR, EAA ve SA için ROC eğrileri**

EAA: esnek ayırma analizi, LR: lojistik regresyon analizi, SA: sinir ağları

## Kaynaklar

- Jo I, Ahn Y, Lee J, Shin KR, Lee HK, Shin C. Prevalence, awareness, treatment, control and risk factors of hypertension in Korea: The Ansan Study. *J Hypertens* 2001; 19: 1523-32.
- Cheitlin MD. Systemic hypertension. In: Sokolow M, McIlroy MB, editors. 6th ed. *Clinical Cardiology*. Norwalk, Connecticut: Appleton & Lange; 1993. p.247-51.
- Hall WD, Wollam G, Tuttle EP Jr. Diagnostic evaluation of the patient with systemic arterial hypertension: An overview. In: Schlant RC, Alexander RW, O'Rourke RA, Roberts R, Sonnenblick E, editors. 8th ed. *Hurst's the Heart*. New York: McGraw-Hill Inc.; 1997. p.1403-25.
- Cruickshank JM, Neil-Dwyer G, Dorrance DE, Hayes Y, Patel S. Acute effects of smoking on blood pressure and cerebral blood flow. *J Hum Hypertens* 1989; 3: 443-9.
- Williams RR. Will gene markers predict hypertension? *Hypertension* 1989; 14: 610-3.
- Williams RR, Hunt SC, Hasstedt SJ, et al. Current knowledge regarding the genetics of human hypertension. *J Hypertens* 1989; 7(Suppl): S8-13.
- Speers MA, Kasl SV, Freeman DH Jr, Ostfeld AM. Blood pressure concordance between spouses. *Am J Epidemiol* 1986; 123:818-29.
- Elfor J, Phillips A, Thomson AG, Shaper AG. Migration and geographic variations in blood pressure in Britain. *BMJ* 1990; 300: 291-5.
- Folkow R. The pathophysiology of hypertension. Differences between young and elderly patients. *Drugs* 1993; 46 (suppl 2): 3-7.
- Kaplan NM. *Clinical hypertension*. Baltimore: Williams and Wilkins; 1994.
- Schieken RM. Genetic factors that predispose the child to develop hypertension. *Pediatr Clin North Am* 1992; 40: 1-11.
- Hosmer DW, Lemeshow S. *Applied logistic regression*. New York: John Wiley & Sons, 2000.
- Özdamar K. *Paket programlar ile istatistiksel veri analizi 1*. Eskişehir: Kaan Kitabevi; 1999.
- Sharma S. *Applied multivariate techniques*. New York: John Wiley & Sons; 1996.
- Özdamar K. *Paket programlar ile istatistiksel veri analizi 2*. Eskişehir: Kaan Kitabevi; 2004.
- Hastie T, Tibshirani R, Buja A. Flexible discriminant and mixture models. In: Kay J & Titterton D, editors. *Proceedings of "Neural Networks and Statistics"* conference, Edinburgh: Oxford University Press; 1995. p.1-23.
- Hastie T, Tibshirani R, Buja A. Flexible discriminant analysis by optimal scoring. *J Am Stat Assoc* 1994; 89: 1255-70.
- Hastie T, Tibshirani R. Discriminant analysis by Gaussian mixtures. *J Roy Stat Soc B* 1996; 58, 155-76.
- Kim H, Loh W-Y. Classification trees with bivariate linear discriminant node models. *J Comput Graph Stat* 2003; 12: 512-30.
- Edler L, Grassmann J, Suhai S. Role and results of statistical methods in protein fold class prediction. *Math Comput Model* 2001; 33: 1401-7.
- Lee HKH. Model selection for neural network classification. *J Classif* 2001; 18: 227-43.
- Abdul-Kareem S, Baba S, Zubairi YZ. Back propagation neural network for medical prognosis: a comparison of different training algorithms. 2001. Available from: [www.sat.ait.ac.th/ejat/articles/3.1/main.html](http://www.sat.ait.ac.th/ejat/articles/3.1/main.html).
- Fine TL. *Feed forward neural network methodology*. New York: Springer; 1999.
- Haykin S. *Neural network: a comprehensive foundation*. New York: Prentice Hall; 1999.
- Hassoun MH. *Fundamentals of artificial neural networks*. Cambridge: Mass MIT Press; 1995.
- Demuth H, Beale M. *Neural network toolbox user's guide*. USA: The Mathworks Inc; 2001.
- Francis L. *The basics of neural networks demystified*. Contingencies 2001: 56-61. Available from: [www.contingencies.org/novdec01/workshop.pdf](http://www.contingencies.org/novdec01/workshop.pdf).
- Dirican A. Evaluation of the diagnostic test's performance and their comparisons. *Cerrahpaşa J Med* 2001; 32: 25-30.
- Hajian-Tilaki KO, Hanley JA, Joseph L, Collet J-P. A comparison of parametric and nonparametric approaches to ROC analysis of quantitative diagnostic tests. *Med Decis Making* 1997; 17: 97-102.
- Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148: 839-43.
- Principe J, Euliano NR, Lefebvre WC. *Neural and adaptive systems: fundamentals through simulations*. New York: John Wiley & Sons Inc; 1999.
- Lee DH, Ha MH, Kim JR, Jacobs DR. Effects of smoking cessation on changes in blood pressure and incidence of hypertension. *Hypertension* 2001; 37: 194-8.
- Harrap SB. Genetics. In: Oparil S, Weber MA, editors. *Hypertension*. Philadelphia: W.B. Saunders Company; 2000. p.29-40.
- Turner ST, Boerwinkle E. Genetics of hypertension, target-organ complications, and response to therapy. *Circulation* 2000; 102 (Suppl): IV40-5.
- Nazlı N. Hipertansiyon konusunda hastaların bilmesi gerekenler ve doğru kan basıncı ölçülmesi. 2003. Available from: [www.hacettepem.org/makaleler.php?a=&b=27 &mNo=47](http://www.hacettepem.org/makaleler.php?a=&b=27 &mNo=47).
- Lemne C, Hamsten A, Karpe F, Ehle PN, De Faire U. Dyslipoproteinemic changes in borderline hypertension. *Hypertension* 1994; 24: 605-10.
- Steinmetz A, Kirklies A, Schlosser G, et al. Lipoprotein (a), low density lipoprotein, intermediate density lipoprotein, and blood pressure in a young male population. *Clin Investig* 1993; 71: 145-9.
- Landsberg L. Obesity-related hypertension as a metabolic disorder. In: Oparil S, Weber MA, editors. *Hypertension*. Philadelphia: W.B. Saunders Company; 2000. p.118-24.
- Breslow J. Genetics of lipoprotein disorders. *Circulation* 1993; 87: 16-21.
- Scanu A, Lawn R, Berg K. Lipoprotein (a) and atherosclerosis. *Ann Intern Med* 1991; 115: 209-18.