

**T. C.
ULUDAĞ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI**

**LOJİSTİK REGRESYON ANALİZİ VE YAPAY SİNİR AĞLARI TEKNİKLERİNİN
SINIFLAMA ÖZELLİKLERİNİN KARŞILAŞTIRILMASI VE BİR UYGULAMA**

Gökhan OCAKOĞLU

(YÜKSEK LİSANS TEZİ)

Bursa-2006



T. C.
ULUDAĞ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

LOJİSTİK REGRESYON ANALİZİ VE YAPAY SİNİR AĞLARI TEKNİKLERİNİN
SINIFLAMA ÖZELLİKLERİNİN KARŞILAŞTIRILMASI VE BİR UYGULAMA

Gökhan OCAKOĞLU

(YÜKSEK LİSANS TEZİ)

Danışman: Prof. Dr. İsmet KAN

Bursa-2006

Kabul Ve Onay Yazısı

Sağlık Bilimleri Enstitüsü Müdürlüğüne,

Bu tez jürimiz tarafından
tezi olarak kabul edilmiştir.

Adı ve Soyadı
İmza

Tez Danışmanı

Üye

Üye

Üye

Üye

Bu tez, Enstitü Yönetim Kurulunun.....tarih,
.....sayılı toplantısında alınan numaralı kararı ile kabul edilmiştir.

.....

Enstitü Müdürü

İÇİNDEKİLER

İÇİNDEKİLER

| | |
|------------------------|-----|
| TÜRKÇE ÖZET..... | II |
| İNGİLİZCE ÖZET..... | III |
| GİRİŞ..... | 1 |
| GENEL BİLGİLER..... | 3 |
| GEREÇ ve YÖNTEM..... | 38 |
| BULGULAR..... | 39 |
| TARTIŞMA ve SONUÇ..... | 47 |
| KAYNAKLAR..... | 50 |
| TEŞEKKÜR..... | 53 |
| ÖZGEÇMİŞ..... | 54 |

TÜRKÇE ÖZET

Bu çalışma, lojistik regresyon analizi ve yapay sinir ağlarının sınıflama etkinliklerini karşılaştırmayı amaçlamaktadır. Lojistik regresyon analizi ve yapay sinir ağları teknikleri, bireylerin sınıflandırma oranlarına göre karşılaştırılmışlardır.

Çalışmaya dahil edilen veri seti, Ercan ve arkadaşları (1) tarafından yapılan çalışmanın veri setinden lojistik regresyon analizi ve yapay sinir ağı tekniklerine uyacak şekilde seçilen 140 klinik hastasından oluşmaktadır. Yapılan analizler sonucunda, örnek veri seti doğru sınıflandırma oranları, lojistik regresyon analizi için % 81,4 ve yapay sinir ağları tekniği için de % 85 olarak hesaplanmış ve çalışmaya alınan veri seti için yapay sinir ağları tekniğinin lojistik regresyona göre daha iyi bir “doğru sınıflandırma oranına” sahip olduğu görülmüştür.

Anahtar Kelimeler: Lojistik regresyon analizi, yapay sinir ağları tekniği, doğru sınıflandırma oranı.

SUMMARY

LOGISTIC REGRESSION ANALYSIS AND COMPARISON OF CLASSIFICATION CHARACTERISTICS OF ARTIFICIAL NEURAL NETWORK TECHNIQUES AND AN APPLICATION

This study was aimed to compare the classification effectivities of logistic regression analysis and artificial neural network. Comparison of logistic regression analysis and artificial neural network techniques was carried out according to individual's classification ratios.

Data set included in the study was selected from the data set of the study done by Ercan et al(1). Data of 140 clinical patients that were appropriate for logistic regression analysis and artificial neural network techniques were included. As a result of the analysis, correct classification ratios of the sample data set for logistic regression analysis and artificial neural network techniques were calculated as 81.4% and 85% respectively. For the data set included, artificial neural network technique was found to have a better "correct classification ratio" than the logistic regression analysis.

Key words: Logistic regression analysis, artificial neural network technique, correct classification ratio.

GİRİŞ

Günlük yaşantımızda çoğu zaman çevremizdeki olayları, nesnelere farkında olmadan, bazı özelliklerini dikkate alarak, sınıflandırarak genelleriz; genelleme sonucunda, sorunlarımıza daha etkin bir biçimde yaklaşım çözümleri üretiriz.

Sınıflama, günlük yaşantımızda olduğu gibi bilimsel çalışmalarda da sorunların çözümünde sağladığı fayda nedeniyle oldukça sık başvurulan bir işlemdir. Tıp alanında, hastalıkların sınıflandırılması ve bu sınıflandırmaya göre tedavi yöntemlerinin geliştirilmesi en belirgin örneklerdendir. Tıbbın yanında diğer bilim dallarında da sınıflandırmanın işlevi görülebilmektedir.

Özellikle tıp ve biyoloji alanında yapılan çalışmalarda, veri setleri oldukça karmaşık bir yapı teşkil etmektedir. Veriler üzerinde çalışılmadan önce, bu karmaşık yapının düzenlenmesi gerekmektedir; düzenleme, belirlenen amaç doğrultusunda sınıflandırmanın yapılması ile mümkün olmaktadır. (1).

Klinik tıbbın amacı, hastaları olabildiği kadar etkin biçimde tedavi etmektir. Bunu başarabilmek için hastaların ıstırap çektiği oluşumu mümkün olduğu kadar doğru olarak teşhis etmek gerekir (2). Teşhis, semptomlara ve klinik bulgulara dayanarak, önce hastalığın sonra da kişinin o hastalık içindeki konumunun belirlenmesiyle konulur. Her hastalığın kendine özgü semptomları vardır; ancak, aynı hastalığın görüldüğü kişilerde hastalığın seyri ve sonucu aynı olmayabilir. Aynı hastalık sınıfı içinde aynı semptomları gösteren kişilerde de bu semptomların şiddeti farklı olabilir (3). Bu nedenle aynı hastalığın görüldüğü kişilerde, eğer var ise, farklılık oluşturacak bir veya birden fazla nitelik açısından belli bir sınıflandırma içinde incelenmesi tedavideki etkinliği şüphesiz artıracaktır.

Belirlenen amaç doğrultusunda birimlerden elde edilen bilgiler yığımından daha fazla fayda sağlanabilmesinde, sınıflandırmanın payı büyüktür. Bu nedenle, belli bir hastalığın görüldüğü olguların, klinik ve epidemiyolojik özellikleri doğrultusunda, sınıflandırılmasında fayda vardır (4, 5).

Bilimsel, çalışmalarda incelenen olaylar birden fazla etkenin etkisi altındadır. Bu nedenle, olayların açıklanmasında verilecek kararlar için birimlerin tek değişkenini veya değişkenlerinden her birini sırayla tek tek göz önüne alarak çıkarsamalar yapmak yerine, birden fazla değişkeni eş zamanlı olarak göz önünde bulundurup çıkarsamalar yapmak daha güvenilir olacaktır. Çünkü, çok değişkenli istatistiksel çözümler genellikle kendi

doğal çevrelerinde bir bütün olarak gözlenirler ve bütünlüğü sağlayan değişkenlerin bağımlılık yapısını açıklamaya çalışırlar (6). Regresyon metodları, bağımlı değişkeniyle bir ya da daha fazla bağımsız değişken arasındaki ilişkiyi tanımlamaya yönelik veri analizlerinin önemli bir parçası olmaya başlamıştır. Modelleme örneklerinin en yaygın olanları bağımlı değişkeninin sürekli olduğu doğrusal regresyon modelleri olsa da, son yirmi yıldır, bağımlı değişkenin binom dağılımına uyduğu durumlarda çok değişkenli lojistik regresyon modelinin kullanımı özel bir araç olmaktan çok, epidemiyolojik çalışmalar için standart bir metot olmuştur (7, 8). Çoklu lojistik fonksiyon ilk olarak 1960'lı yıllarda Cornfield ve arkadaşları tarafından bağımsız değişkenlerin bileşik dağılımı hem hasta, hem de kontrol grubu için aynı varyans-kovaryans matrisiyle çok değişkenli normal dağılıma uyduğunda risk kestiricisi olarak kullanılmıştır (9).

LRA'nın kullanım amacı: En az değişkeni kullanarak en iyi uyuma sahip olacak şekilde bağımlı ile bağımsız değişkenler kümesi arasındaki ilişkiyi tanımlayabilen ve istatistiksel olarak kabul edilebilir modeli kurmaktır.

Regresyon modellerinin çözüm yöntemlerinden biri de son yıllarda kullanılmaya başlanan sinir ağları yaklaşımıdır. İlk yapay sinir ağı modeli 1943 yılında Warren McCulloch ve Walter Pitts tarafından gerçekleştirilmiştir. 1980'li yıllar ise yapay sinir ağları (YSA) ile ilgili çalışmalar için bir atılım dönemi olmuştur.

YSA bilgi sınıflama ve bilgi yorumlamanın da içinde bulunduğu çok değişik problemlerin çözümünde kullanılmanın yanı sıra sağlık, iş hayatı, finans, endüstri ve eğitim alanlarında var olan yöntemlerin yerine başarıyla uygulanmaktadır (10). 1990'lı yılların başlangıcında yayınlanan bazı kaynaklarda bazı YSA modelleri ile bazı istatistik tekniklerin benzer hatta bazılarının aynı olduğuna dikkat çekilmiştir. Sonraki çalışmalar ise bunun tesadüfi olmadığını, bu iki alanın birbirleriyle ileri derecede ilişkili olduğunu göstermiştir. Yapay sinir ağı modelleri ile istatistik tekniklerin karşılaştırılması birinin diğerinin geliştirilmesinde önemli olduğunu ortaya çıkarmıştır. Perseptron, çok katmanlı perseptron gibi bazı yapay sinir ağı modelleri istatistiksel uygulamalar için faydalı olabileceği ve aynı şekilde tahmin ölçütü, güven aralıkları, diagnostik yöntemler gibi bazı istatistiksel tekniklerin de yapay sinir ağı uygulamalarına uygulanabileceği bazı bilim adamları tarafından ileri sürülmektedir. YSA ve istatistiksel metodoloji arasındaki iletişimin geliştirilmesi her iki alan için de büyük yarar sağlamaktadır (11).

Bu çalışmada, LRA ve YSA tekniklerinin sınıflama özellikleri karşılaştırılarak, gerçek bir veri seti üzerinde uygulama yapılmış ve iki yöntemin sınıflama özellikleri tartışılmıştır.

GENEL BİLGİLER

Yapay Sinir Ağı Nedir?

Bir yapay sinir ağı, biyolojik sinir ağlarının karakteristiklerine benzer karakteristiklere sahip bir bilgi işleme sistemidir. YSA, insanın idrak etmesi ve biyolojik nöron yapısının matematiksel modelinin aşağıdaki kurallar varsayılarak genelleştirilmesi sonucunda oluşturulmuştur:

- Bilgi işleme, nöron adı verilen birçok basit elemanlarda gerçekleşir;
- Sinyaller, nöronlar arasındaki ilişkiyi sağlayan bağlantılarla iletilir;
- Her bir bağlantının bir ağırlık değeri vardır ve bu değer, gerçek nöronlarda olduğu gibi sinyal geçişini üretmektedir;
- Sinir ağı içindeki her bir nörona aynı bir aktivasyon fonksiyonu uygulanır (genelde bu doğrusal olmayan bir fonksiyondur) ve bu fonksiyonun çıkış değeri sayesinde nöronun çıkış sinyali hesaplanır;

Herhangi bir yapay sinir ağı;

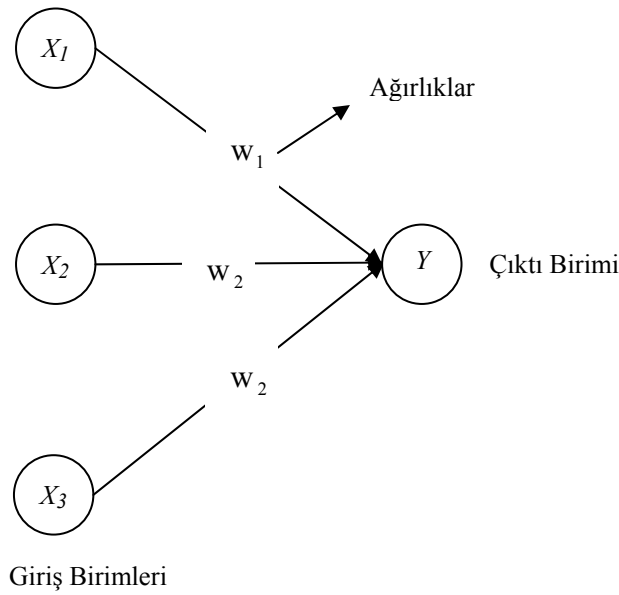
- Nöronlar arasındaki bağlantının bir modeli yani mimarisi ile,
- Bağlantılardaki ağırlıkların hesaplanması (bu hesaplama, eğitim kuralı ya da öğrenme algoritması olarak da adlandırılır) ile
- Aktivasyon fonksiyonu ile tanımlanabilir (12).

Bir yapay sinir ağı, nöron, birim, hücre ya da düğüm olarak adlandırılan çok sayıda basit işlem birimlerinden oluşur. Her bir nöron, diğer bir nörona belli bir ağırlık değerine sahip olan haberleşme bağlantılarıyla bağlanır. Ağırlıklar, yapay sinir ağının bir problemi çözmesi için gerekli olan bilgiyi hazırlamaktadır. YSA çok çeşitli problemlerin çözümünde kullanılabilirler. Örnek olarak, bilgileri ve numuneleri saklamada ve onları daha sonra tekrar tanımada, numuneleri sınıflandırmada, giriş numunelerinin çıkış numunelerine dönüştürülmesinde, benzer örneklerin gruplandırılmasında ya da doğal olmayan optimizasyon problemlerinin çözümlerinin bulunmasında ve daha pek çok alanda YSA çok geniş bir biçimde kullanılabilir.

Her nöronun bir iç durumu vardır ve bu iç durum aktivasyon ya da aktivasyon düzeyi olarak adlandırılır. Bu düzey, alınan giriş değerlerinin bir fonksiyonudur. Herhangi bir nöron, kendi aktivasyonunun, genelde sinyal şeklinde diğer nöronlara gönderir. Ama bu sinyal birden fazla nörona aynı anda gönderilebilir.

Örnek olarak şekil-1’de gösterilen bir Y nöronunu düşünelim. Bu nöron X_1, X_2 ve X_3 nöronlarından giriş sinyallerini alır. Bu nöronların aktivasyonları yani çıkış sinyalleri, sırasıyla x_1, x_2 ve x_3 ’tür. Bağlantılar üzerindeki ağırlıklar X_1, X_2 ve X_3 nöronlarından Y nöronuna doğru sırasıyla w_1, w_2 ve w_3 ’tür. Ağ girişi olan y_{in} değeri X_1, X_2 ve X_3 ’den Y’ye giden ağırlıklı sinyallerin toplamıdır. y_{in} değeri 1. eşitlikteki gibi hesaplanır.

$$y_{in} = w_1x_1 + w_2x_2 + w_3x_3 \quad 1.$$



Şekil-1 Basit bir yapay nöron

Y nöronunun aktivasyonu y , ağa giriş değerlerinin bir fonksiyonu olarak tanımlanır.

$$y = f(y_{in}) \quad 2.$$

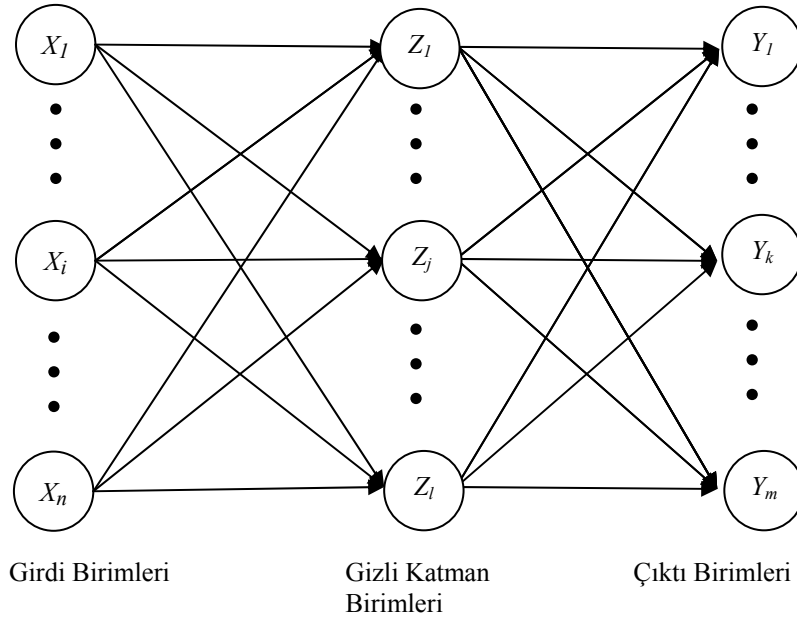
Bu fonksiyon, S-biçimli sigmoid fonksiyon olabilir. Sigmoid fonksiyonun formu aşağıda verildiği gibidir.

$$f(x) = \frac{1}{1 + \exp(-x)}$$

Bu fonksiyonu, diğer aktivasyon fonksiyonlarından biri de olabilir.

Y nöronunun Z_1 ve Z_2 nöronlarına v_1 ve v_2 ağırlıklarıyla bağlandığı varsayalım. Bu durum, şekil-2’de gösterilmektedir. Y nöronu y sinyalini diğer birimlere gönderir. Bununla birlikte, genel olarak Z_1 ve Z_2 nöronları tarafından alınan sinyaller farklı olmaktadır; çünkü her bir sinyal aktarıldığı bağlantıda bulunan v_1 ve v_2 ağırlıkları ile orantılıdır. Z_1 ve Z_2 ’nin aktivasyonları olan z_1 ve z_2 değerleri, sadece tek bir nörona bağlı değildir. Onlar birbirinden farklı birden fazla nörondan gelen sinyallere bağlıdır.

Şekil-2’deki yapay sinir ağı basit olmasına rağmen, gizli birimin görünümü ve doğrusal olmayan aktivasyon fonksiyonu sayesinde birçok problemi çözebilir. Başka bir yönden gizli bir birime sahip yapay sinir ağının öğretilmesi yani ağırlıkların optimal değerlerinin bulunması oldukça zordur.



Şekil-2 Basit bir yapay sinir ağı

Biyolojik Sinir Ağları

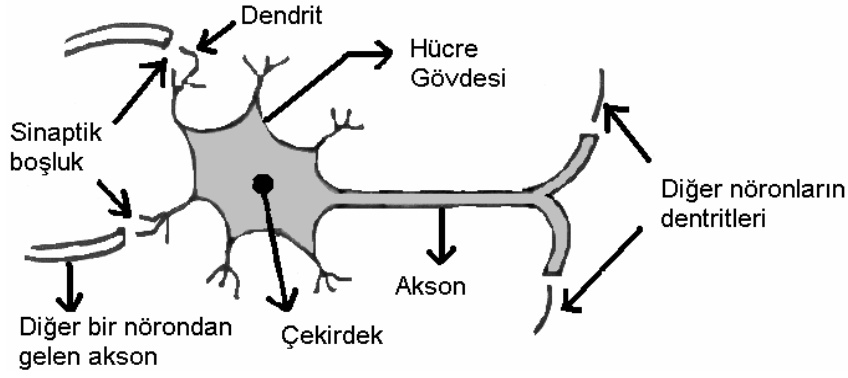
Bir biyolojik nöronun üç tip bileşeni vardır. Bu bileşenler, dentritler, soma ve aksondur. Birçok dentrit, sinyalleri diğer nöronlardan alır. Alınan bu sinyaller, dentritler arasındaki sinaptik boşluklardan iletilen elektrik tepkileridir. Kimyasal vericilerin hareketi, giren sinyalleri değiştirmektedir. Bunlar, bir anlamda YAS daki ağırlıklara benzemektedir (13).

Biyolojik bir nöronun yapısı şekil-3'de gösterilmiştir. Yapay sinir ağı işlem elemanlarının birkaç önemli özelliği, biyolojik nöronların özelliklerini akla getirmektedir. Bu özellikler aşağıdaki gibidir:

- Her bir işlem elemanı yani her bir nöron, farklı sinyaller alabilir;
- Alınan sinyaller sinaptaki ağırlıklar tarafından değiştirilebilir;
- İşlem elemanları yani nöronlar, ağırlıklı girdileri toplar;
- Uygun koşullar altında yani yeterli girdi sağlandığında, işlem elemanı çıkış sinyalini aktarır;
- Belirli bir nöronun çıkışı diğer birçok nörona gidebilir.

YSA'nın biyolojik nöronlardan aldığı önemli diğer özellikler aşağıda verilmiştir:

- Bilgi işleme yereldir;
- Bellek ikiye ayrılmıştır:
 - a) Uzun dönemli bellek, nöronun sinapslarındaki ağırlık değerlerine bağımlı olur
 - b) Kısa dönemli bellek, nöronlar tarafından gönderilmiş daha hızlı sinyallere uymaktadır.
 1. Bir sinapsın gücü deneyimlerle değiştirilebilir;
 2. Nöro aktarıcılar, sinapslar için hızlandırıcı ya da yavaşlatıcı olabilir.

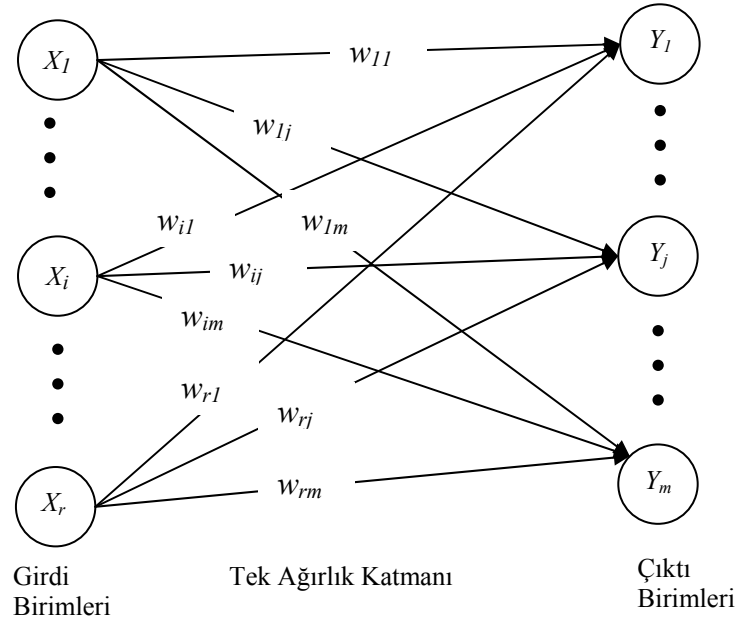


Şekil-3 Biyolojik nöronun yapısı

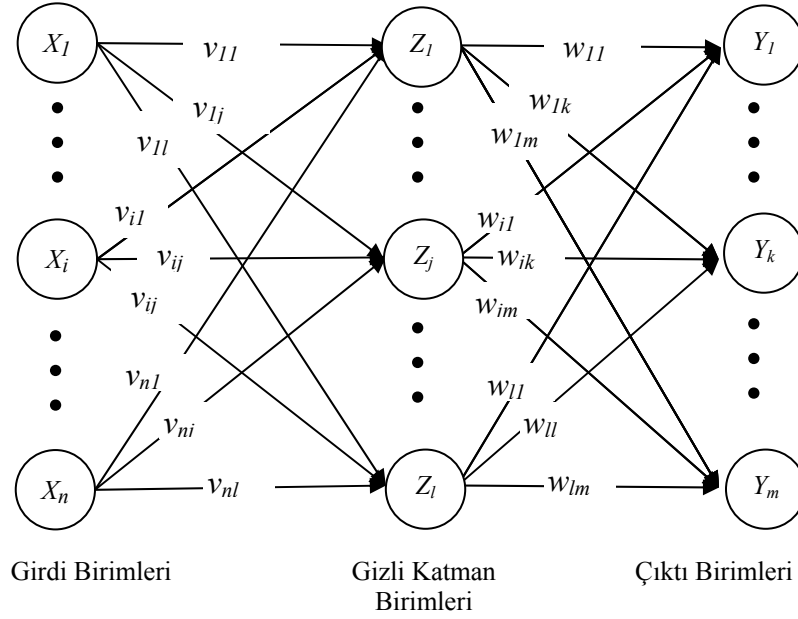
YSA' nın biyolojik sinir sistemiyle paylaştığı diğer bir karakteristik de hata toleransıdır. Biyolojik nöral sisteminde iki hata toleransı olabilir. Birincisinde, giriş sinyallerinde daha önce hiç görülmemiş bir sinyalle karşılaşılabilir. Buna bir örnek olarak, daha önceden hiç görülmemiş ya da çok uzun süre önce görülmüş bir insanı, bir fotoğrafın içinde hemen tanıma yeteneği verilebilir. İkincisinde, sinir sisteminin kendisine zarar vermesine izin verilebilir. İnsan 100 milyardan fazla nöronla dünyaya gelir. Nöronların büyük bir kısmı beyindedir ve bu nöronlar öldüğü takdirde, onların birçoğu artık yenilenmez. İnsan beyni birçok nöronunu kaybetmesine rağmen öğrenmesine devam etmektedir. Travmatik nöron kayıplarında bile, diğer nöronlar, ölen nöronların fonksiyonlarını yerine getirmek için eğitilebilir. Benzer şekilde, YSA da, ağdaki küçük hasarlara karşı duyarsız şekilde tasarlanabilir ve sinir ağının önemli hasarlar sonucunda dahi tekrar eğitilebilir.

Tipik Mimarileri

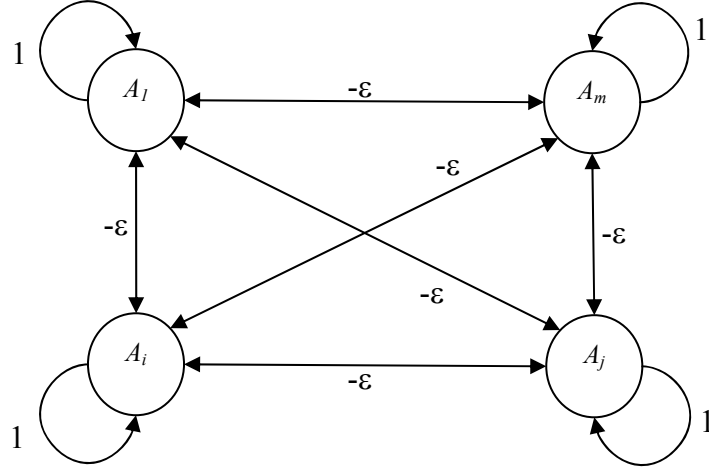
Sinir ağları, tek katmanlı ya da çok katmanlı olarak sınıflandırılırlar. Katman sayısını belirlerken, girdi birimi bir katman olarak sayılmaz; çünkü bunlar üzerinde hiçbir hesaplama işlemi yapılmaz. Bir ağ içindeki katman sayısı, nöronları bağlayan ağırlıklı bağlantı sayısına eşittir. Örneğin şekil-2' de gösterilen sinir ağının, iki ağırlıklı katmanına sahiptir. Tek katmanlı ve çok katmanlı sinir ağlarının farklı iki örneği, şekil-4 ve şekil-5' de gösterilmiştir. Bu ağlar ileri beslemeli ağlara birer örnektir. Bu çeşit ağlarda girdi sinyalleri girdi biriminden çıktı birimine doğru ilerlemektedir. Şekil-6' da geri dönüşümlü sinir ağına bir örnek verilmiştir. Bu ağda kapalı bir sinyal döngüsü vardır.



Şekil-4 Tek katmanlı bir yapay sinir ağı



Şekil-5 İki katmanlı bir yapay sinir ağı



Şekil-6 Rekabetçi katmana sahip bir yapay sinir ağı

Tek ve Çok Katmanlı Sinir Ağları

Tek katmanlı yapay sinir ağında bir tane ağırlıklı bağlantı katmanı bulunur. Birimler çoğu kez, sinyalleri alan girdi birimi ve ağı cevabının alınacağı çıktı birimi olmak üzere ikiye ayrılır. Tipik bir tek katmanlı sinir, şekil-4’ de gösterilmiştir. Bu şekilde, girdi birimlerinin tamamı çıktı birimlerine bağlanmıştır.

Çıktı birimleri ile gizli birimler arasında bir ya da daha çok katman olan sinir ağına *çok katmanlı ağ* adı verilir. Tipik olarak, gizli birimlerle ve çıktı birimleri arasında ağırlıklı bağlantı katmanı bulunan ağ mimarisidir. Çok katmanlı sinir ağları tek katmanlıya göre çok daha karmaşık problemleri çözebilir; fakat eğitimleri çok daha zor olabilir. Tek katmanlı bazı ağların problemi tam olarak çözecek şekilde eğitilemediği durumlarda son derece başarılı olabilirler (12).

Ağırlıkların Ayarlanması

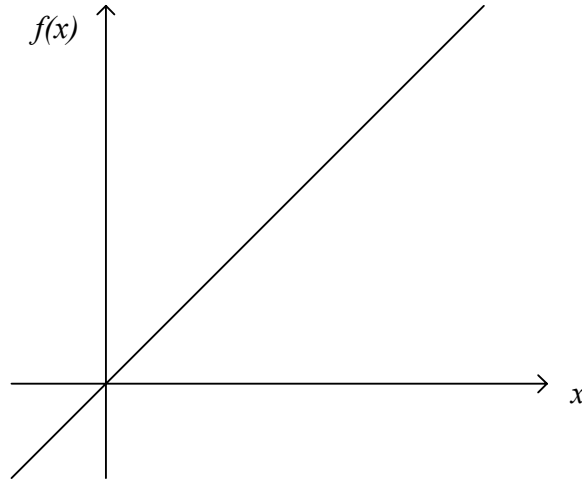
Farklı sinir ağı karakteristiklerini birbirinden ayıran en önemli etken, onların eğitilmeleri esnasında bağlantı ağırlıkları değerlerinin ayarlanmasıdır. İki tür öğrenme çeşidi birbirinden ayırt edilebilir: kontrollü ve kontrolsüz öğrenme.

Aktivasyon Fonksiyonları

Daha önceden de bahsedildiği gibi, bir yapay sinir ağında yapılması gereken en temel işlemler, ağırlıklı girdi değerlerinin toplanması ve bir çıktı ya da aktivasyon fonksiyonu uygulamasıdır. Aktivasyon fonksiyonlarının çeşitleri aşağıdaki gibi verilebilir.

Özdeşlik Fonksiyonu

Tüm x ler için $f(x) = x$ dir.



Şekil-7. Özdeşlik fonksiyonu

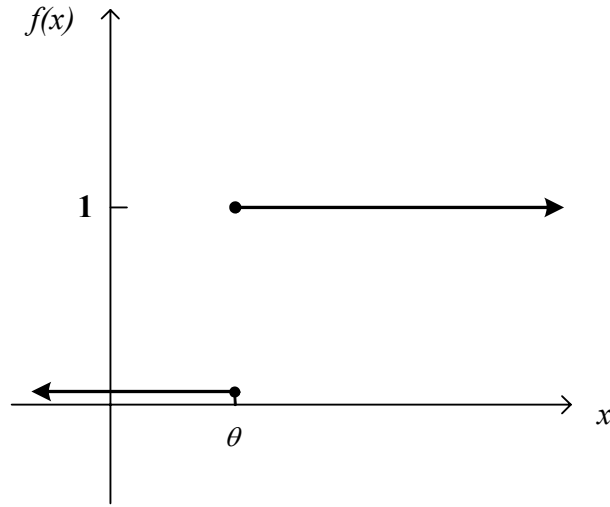
Belirli bir tabakadaki tüm nöronlar için aynı aktivasyon fonksiyonu kullanılır. Birçok durumda doğrusal olmayan aktivasyon fonksiyonları kullanılır. Tek katmanlı ağlara göre çok katmanlı ağların avantajlarının kullanabilmek için, doğrusal olmayan aktivasyon fonksiyonlarının kullanılması gereklidir. Tek katmanlı ağlar, genelde, adım fonksiyonunun giriş değerlerini, diğeri ikili (0 ya da 1 veya -1 ya da 1) olan çıkış değerline çevirmek için kullanılır (Şekil-8).

Θ Eşik Değeri ile Verilen İkili Adım Fonksiyonu

Bu fonksiyon, aşağıdaki formülle verilebilir:

$$f(x) = \begin{cases} 0 & , x < \theta \text{ ise} \\ 1 & , x \geq \theta \text{ ise} \end{cases} \quad 4.$$

İkili adım fonksiyonu eşik fonksiyonu ya da Heaviside fonksiyonu olarak da bilinmektedir. Aşağıdaki ikili adım fonksiyonu görülmektedir:



Şekil-8. İkili adım fonksiyonu

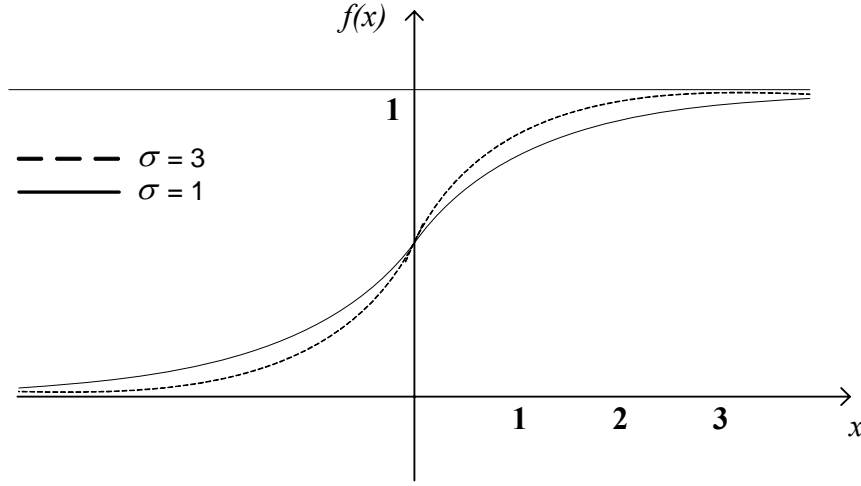
Sigmoid fonksiyonları, oldukça kullanışlı aktivasyon fonksiyonlarıdır. Sigmoid fonksiyonları içinde *lojistik* ve *hiperbolik tanjant* fonksiyonları en yaygın olanlarıdır. Özellikle geriye yayılım öğrenme algoritmasında bu fonksiyonların kullanımı avantajlıdır. Çünkü fonksiyonun belirli bir noktadaki değeri ile onun türevinin değeri arasındaki ilişki öğrenme zamanındaki hesap yükünü azaltır. Lojistik fonksiyon, değerleri 0 ile 1 arasında değişen bir sigmoid fonksiyonudur ve sinir ağları için aktivasyon fonksiyonu olarak çoğunlukla kullanılır (12). Fonksiyonun aralık değerini vurgulamak için ikili sigmoid adı verilir. Bu fonksiyona *lojistik sigmoid* de denir. Şekil-9’ da, adım parametresi σ nın farklı değerleri için *ikil sigmoid* eğrileri gösterilmiştir.

İkili Sigmoid Fonksiyonu

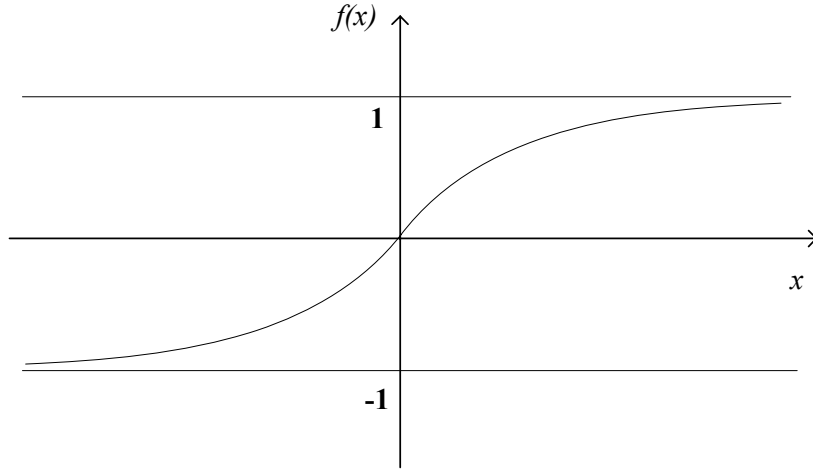
$$f(x) = \frac{1}{1 + \exp(-\sigma x)} \quad 5.$$

$$f'(x) = \sigma f(x)[1 - f(x)] \quad 6.$$

Lojistik sigmoid fonksiyonu istenen deęer aralıęına göre ölçeklenebilir ve böylece probleme uygun bir fonksiyon haline gelebilir. En yaygın aralık, -1 ile 1 arasındadır. Bu sigmoid fonksiyonu, bipolar sigmoid olarak adlandırılır. Şekil-10' da $\sigma = 1$ için böyle bir fonksiyon gösterilmiştir:



Şekil-9 İki kutuplu sigmoid fonksiyonu (adım parametreleri $\sigma = 1$ ve $\sigma = 3$ için)



Şekil-10 İki kutuplu sigmoid fonksiyonu ($\sigma = 1$ için)

İki Kutuplu Sigmoid Fonksiyonu

$$g(x) = 2f(x) - 1 = \frac{2}{1 + \exp(-\sigma x)} - 1$$

$$= \frac{1 - \exp(-\sigma x)}{1 + \exp(-\sigma x)} \quad 7.$$

$$g'(x) = \frac{\sigma}{2} [1 + g(x)][1 - g(x)] \quad 8.$$

Bipolar sigmoid hiperbolik tanjant fonksiyonuna çok yakındır. Hiperbolik tanjant fonksiyonunda istenilen çıkış aralığı (-1 ile +1) arasında ise aktivasyon fonksiyonu olarak kullanılmaktadır. İkisi arasındaki benzerlik $\sigma = 1$ için gösterilecek olursa:

$$g(x) = \frac{1 - \exp(-x)}{1 + \exp(-x)} \quad 9.$$

Hiperbolik tanjant fonksiyonu ise

$$h(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} = \frac{1 - \exp(-2x)}{1 + \exp(-2x)} \quad 10.$$

dir.

Hiperbolik tanjant fonksiyonunun türevi :

$$h'(x) = [1 + h(x)][1 - h(x)] \quad 11.$$

biçiminde elde edilir.

Eğer ikili veri 0 ile 1 arasında değişiyorsa bunu bipolar biçime sokarak bipolar sigmoid şekline ya da hiperbolik tanjant fonksiyonu şeklinde kullanmak genelde tercih edilen yöntemdir.

Ağ Girişlerinin Hesaplanması İçin Matris Çarpma Metodu

Sinir ağlarının bağlantılı ağırlıklarını $\mathbf{W} = (w_{ij})$ matrisinde kaydedelim. Bu halde, Y_j ' nin ağ girdisi $\mathbf{x} = (x_1, \dots, x_i, \dots, x_n)$ vektörünün $w_{.j}$ ağırlık matrisinin j. kolonu ile (eğer j. elemanda sapma yoksa) basit bir çarpılması işlemi ile gerçekleştirilir :

$$y_{in_j} = x \cdot w_{.j} = \sum_{i=1}^1 x_i \cdot w_{ij} \quad 12.$$

Sapma

Sapma değeri, x vektörüne $x_0 = 1$ değeri eklenerek oluşturulabilir.

$$\mathbf{x} = (1, x_1, \dots, x_i, \dots, x_n) \quad 13.$$

Sapma aynı diğer ağırlıklar gibi davranır, yani $w_{0j} = b_j$ dir. Y_j birimi için ağ girdisi şu şekilde hesaplanır:

$$y_{in_j} = x \cdot w_{.j} = \sum_{i=0}^n x_i \cdot w_{ij} = w_{0j} + \sum_{i=1}^n x_i \cdot w_{ij} = b_j + \sum_{i=1}^n x_i \cdot w_{ij} \quad 14.$$

Bir YSA düğümünün görevi, girişindeki sayıları kendi ağırlık değerleri ile çarpıp, sonra da bu çarpımları toplayıp, toplamı bir yumuşatma fonksiyonundan (genelde *sigmoid* $f(x) = \frac{1}{1+e^{-x}}$ veya *tanh*) geçirdikten sonra çıkışa vermektir. Ancak giriş ve çıkış katmanındaki nöronlar bu kuralın dışındadır. Giriş katmanındaki nöronlar ise sadece kendi girişlerindeki verilerin uygun ağırlıklarla çarpılmış durumlarını toplayıp saklarlar. Bu işleme *ilerleme* denir.

Basit Sinir Ağlarıyla İlgili Genel Bilgiler

Tek katmanlı YSA örnek sınıflandırma, tanıma, örnek ilişkilendirme ve bunun gibi diğer problemlerin çözülmesinde kullanılabilir. Örnek sınıflandırma problemlerinde, her giriş vektörü (örnek, numune) belirli sınıflara ait olabilir ya da olmayabilir. Basit olarak, bir sınıfa üye olma sorusu göz önünde bulundurulur. Çıkış birimi için +1 cevabının alınmasıyla örneğin o sınıfa üye olduğu, -1 cevabı alınırsa, örneğin o sınıfa üye olmadığı belirlenir. Bu tip durumlarda, her bir sınıf için bir çıkış birimi vardır. Örnek tanımlama, örnek tanıma uygulamasının bir çeşididir. Örneklerin ilişkili hatırlanması ise daha farklıdır.

Tek katmanlı sinir ağlarının eğitilmesinde üç önemli yöntem vardır:

- Hebb Kuralı
- Perseptron Öğrenme Kuralı
- Delta Kuralı

Hebb Kuralı

Hebb kuralı, bir yapay sinir ağı için, en eski ve en basit öğrenme kuralı olarak bilinir. Hebb, öğrenmenin, sinaps uzunluklarını (ağırlıkları) değiştirerek meydana geleceğini önermiştir. Hebb'e göre, eğer birbiri ile bağlı iki nöronun her ikisi de aynı zamanda "aktif" ise, bu nöronlara uygun ağırlıkların artırılması gerekmektedir. Benzer olarak, eğer her iki nöron aynı zamanda "pasif" ise, ağırlıkların artırılması gerekir. Bu durumda, daha güçlü bir öğrenme şekli meydana gelir. Geliştirilmiş Hebb kuralı ile tek katmanlı ileri beslemeli bir sinir ağının eğitilmesi bir Hebb ağını anlatır. Hebb kuralı, diğer özgül ağları eğitmek için de kullanılabilir. Tek katmanlı bir yapay sinir ağında birbiri ile bağlantılı nöronlardan bir tanesi giriş birimi, bir tanesi de çıkış birimi olacaktır (hiçbir giriş birimi birbiri ile bağlanmadığı için, herhangi bir çıkış birimleri de birbiri ile bağlı değildir (12, 14).

Perseptron

Perseptronlar, YSA'nın öğrenilebilir niteliğini taşıyan ilk modelidir. Hebb kuralından daha yetenekli bir öğrenme kuralıdır. Perseptron tekrarlı öğrenme algoritmasıdır ve çözümün varlığı durumunda yakınsama niteliğine sahiptir. Bu, perseptron modelinin en önemli niteliklerinden biridir.

Rosenblatt (1962) ve Minsky-Papert (1969, 1988) tarafından çeşitli perseptron modelleri tanımlanmıştır. Orijinal perseptronlar, duyumsal birimler, birleştirici birimler ve cevap birimleri olmak üzere nöronların üç durumuna sahiptirler. Örneğin, bir basit perseptron duyumsal ve birleştirici birimler için ikili aktivasyon, cevap birimi için ise +1, 0, veya -1 değerlerini üreten aktivasyon uygulayabilir.

Sınıflandırma problemlerinde eşik değerli aktivasyon fonksiyonu kullanılır:

$$f(y_{in}) = \begin{cases} -1, & y_{in} < -\theta \text{ ise} \\ 0, & -\theta \leq y_{in} \leq \theta \text{ ise} \\ 1, & y_{in} > \theta \text{ ise} \end{cases}$$

15.

Çıktı biriminin aktivasyonu $y = f(y_{in})$ şeklinde hesaplanır.

Birleştirici birimden cevap birimine giden bağlantıların ağırlıkları perseptron öğrenme kuralı ile ayarlanır. Her eğitim girişi için, sinir ağı, çıkış biriminin cevabını hesaplar. Daha sonra sinir ağı, bu örnek için çıkış değeri ile hedeflenen çıkış arasındaki farkı karşılaştırarak bir hata oluşup oluşmadığını tespit eder. Yapay sinir ağı, hesaplanmış çıkış değeri “0” ve hedef değeri “-1” olan örnek için hatayı ayırt edemez, buna karşılık olarak hesaplanmış çıkış değeri “+1” ve hedef değeri “-1” olan örnek için hatayı ayırt edebilir. Bu durumlarda, hedef verinin işareti yönünde ağırlıkların işareti değiştirilmelidir. Bununla birlikte çıkış birimine “0” olmayan sinyaller gönderen bağlantıların ağırlıkları ayarlanmalıdır. Eğer belirli bir eğitim girişi örneğinde hata oluşuyorsa, ağırlıklar

$$w_i(yeni) = w_i(eski) + \alpha t x_i$$

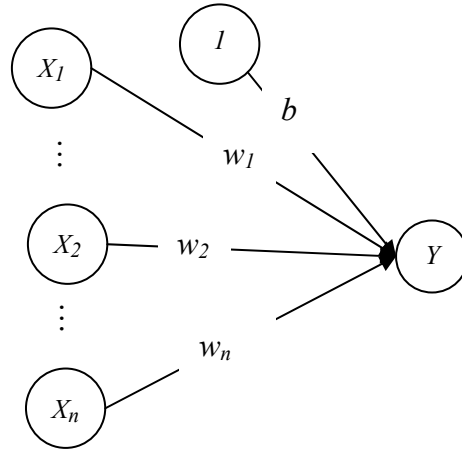
16.

formülüne göre değiştirilmelidir.

Burada hedef değeri t ya “+1” ya da “-1”dir ve α öğrenme oranı katsayısıdır. Eğer hata oluşmadıysa ağırlıklar değiştirilmemelidir. Eğitim işlemi hata oluşmayıncaya kadar devam etmelidir. Perseptron öğrenme kuralı yakınsama teoremine göre, eğer ağda tüm eğitim örnekleri için uygun ağırlıkların varlığına izin verilirse, bu ağırlıklar, eğitim sürecinde bütün eğitim örnekleri için elde edilebilir. Bu kuralın amacı, ağın tam olarak doğru cevap veremediği eğitim örnekleri için ağırlıkları ayarlamaktır. Ayrıca, eğitim sonunda bu ağ sınırsız sayıda eğitim adımları için ağırlıkların değerlerini bulmalıdır (12).

Sınıflandırma Problemleri için Perseptron Algoritması

Şekil-11' de perseptronun mimarisi gösterilmiştir. Burada X_1, \dots, X_n girdi birimleri, Y çıktı birimi ve I sapma sinyalidir. b sapma ağırlığı, w_i ($i = 1, \dots, n$) ağırlıklardır.



Şekil-11 Basit bir perseptron mimarisi

Sınıflandırma problemlerinde, sinir ağının görevi tüm giriş örneklerinin belirli bir sınıfa ait olup olmamasını belirlemektir. Sınıfa ait olma çıkışın “+1” değerine, ait olmama ise çıkışın “-1” değerine uygun olmasıyla belirlenir. Sınıflandırma işlemi yapılabilmesi için ağ, tekrarlı bir teknik ile eğitilir. Girdi ve hedefler ikili ve ya iki kutuplu olabilir. θ eşik değeri tüm birimler için değişmezdir. Sapma ve eşik değerinin her ikisinin aynı zamanda kullanılmasına ihtiyaç duyulmaktadır. Bu işlemin algoritması aşağıda verilmiştir. Bu algoritma, ağırlıkların başlangıç değerlerine ve öğrenme oranına tam olarak duyarlı değildir.

Adım 0 Ağırlıklar ve sapmalara başlangıç değerleri ata.

(Ağırlıkları ve sapma değeri kolaylık için “0” olarak alınabilir.)

Öğrenme oranı olan α ($0 < \alpha \leq 1$)’yü ayarla.

(kolaylık için, α , 1’e eşitlenebilir.)

Adım 1 Durma koşulu yanlış iken , adım 2-6’ yı uygula.

Adım 2 Her bir **s:t** öğrenme çifti için, 3-5 adımlarını uygula.

Adım 3 Giriş birimlerinin aktivasyonlarını ayarla.

$$x_i = s_i \quad i = 1, \dots, n$$

Adım 4 Her çıktı birimi için aktivasyonları hesapla.

$$y_in_j = b_j + \sum_i x_i w_{ij}. \quad j = 1, \dots, m :$$

$$f(y_in) = \begin{cases} -1, & y_in < -\theta \text{ ise} \\ 0, & -\theta \leq y_in \leq \theta \text{ ise} \\ 1, & y_in > \theta \text{ ise} \end{cases}$$

Adım 5 Ağırlıkları ve sapmaları ayarla.

eğer $t_j \neq y_j$ ise,

$$b_j(\text{yeni}) = b_j(\text{eski}) + t_j$$

$$w_{ij}(\text{yeni}) = w_{ij}(\text{eski}) + t_j x_i.$$

eğer $t_j = y_j$ ise,

$$b_j(\text{yeni}) = b_j(\text{eski})$$

$$w_{ij}(\text{yeni}) = w_{ij}(\text{eski})$$

Adım 6 Durma koşulu:

Eğer adım 2'de hiç bir ağırlık değişmezse dur; aksi durumda devam et.

Algoritmada çıktı birimlerinin sayısı $m = 1$ olabilir. Örneğin, mantıksal fonksiyonları gözden geçirirken çıktı biriminin sayısının bir olduğu kabul edilir. Eğitimden sonra, ağ her bir eğitim vektörünü doğru şekilde sınıflandırır.

Sınıflandırma ile ilgili perseptron eğitim algoritmasında, bir ayırma doğrusu yerine, pozitif cevaplar bölgesini sıfır cevaplar bölgesinden ayıran $w_1 x_1 + w_2 x_2 + b > \theta$ doğrusu ve negatif cevaplar bölgesini sıfır cevaplar bölgesinden ayıran $w_1 x_1 + w_2 x_2 + b < -\theta$ doğrusu olmak üzere iki ayırma doğrusu vardır (12).

Delta Kuralı

Delta kuralı, Widrow ve Hoff (1960) tarafından ADALINE için ortaya atılmış olan iteratif bir öğrenme sürecidir. Delta kuralında, tüm girdi numuneleri için çıktı ve hedef farkları karelerinin toplamının, başka bir ifadeyle, toplam hatanın küçültülmesine hedeflenmiştir. Uygun algoritmalarda her numune için gradient vektörünün ters yönünde ağırlıkların güncellenmesi yapılır. Bu durumda delta kuralı, nöron bağlantılarının ağırlıklarını, ağ girişi (y_in) ve ağın hedef çıkışı (t) arasındaki farkı en aza indireyecek şekilde değiştirir. Amaç, tüm eğitim numunelerinin hatalarını en aza indirmektir. Ağırlık

düzeltilmeleri, çok sayıdaki eğitim numunesi ile beraber biriktirilebilir ve bu yığın güncelleştirilmesi olarak adlandırılır.

Çok Katmanlı Yapay Sinir Ağ Modelleri

Tek katmanlı ağlar ayrılamayan problemlerin çözümünde başarısız olmaktadır. Bunun içinde bilim adamları çok katmanlı YSA modellerini incelemişlerdir. Burada önemli aşamalardan biri bu tip ağlar için akıllı bir eğitim algoritmasını geliştirmektir. 1986 yılında Rumelhart, Hinton ve Williams tarafından bu gerçekleştirildi. Benzer fikirler daha önce Werbos (1974), Parker (1985), Cun (1985) gibi bilim adamlarının yayınlamış olduğu makalelerde gözükmektedir. Standart geriye yayılım (backpropagation) olarak adlandırılan bu eğitim metodu hata kareler toplamının geriye yayılım yöntemiyle küçültülmesi fikrine dayanır ve genelleştirilmiş delta kuralını kullanır. Dolayısıyla bu yöntem her adımda hatanın küçültülmesi için, Widrow-Hoff eğitiminde olduğu gibi, gradient azalış yöntemini kullanır. Bu durumda gizli katmanda doğrusal olmayan aktivasyon fonksiyonlar, örneğin lojistik sigmoid fonksiyonu ve ona uygun olarak genelleştirilmiş delta kuralı uygulanmaktadır. Farklı geriye yayılım yöntemlerinin temelinde aşağıdaki aşamalar uygulanır:

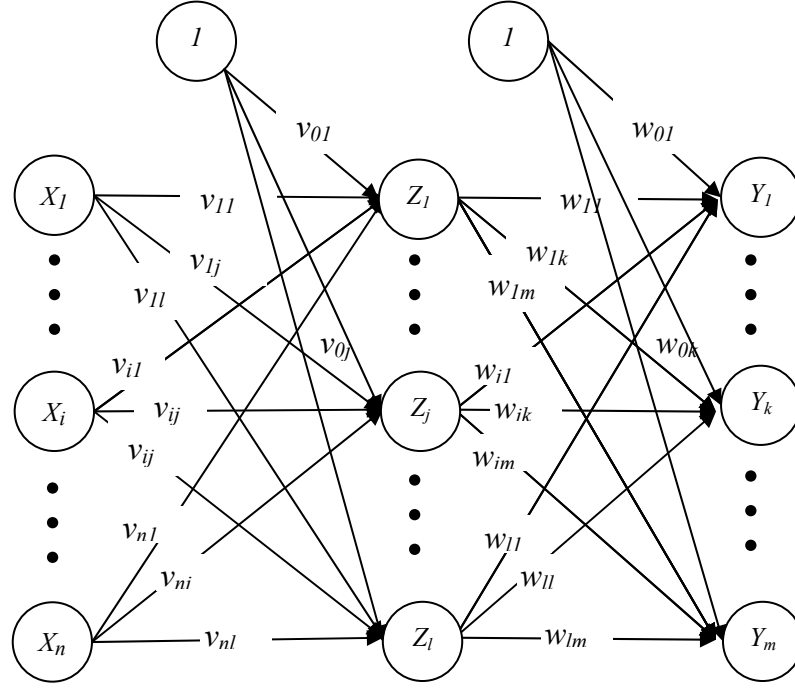
- Girdi eğitim örneklerinin ileri beslemesi,
- Birleşmiş hatanın hesaplanması ve geriye yayılması,
- Ağırlıkların güncellenmesi.

Eğitimden sonra ağ uygulaması sadece ağın ileri besleme aşamasını gerçekleştirir ve hesaplamaları yapar (12).

Standart Geriye Yayılım Ağ Mimarisi

Şekil-12' de bir gizli katmana sahip çok katmanlı ileri beslemeli bir ağ gösterilmiştir. Burada X girdi katmanı, Z gizli katman, Y çıktı katmanına ait birimlerdir. Gizli katmanın j . birimine dahil olan sapma ağırlıkları v_{0j} ($j = 1, \dots, l$), çıktı katmanının k . birimine dahil olan sapma ağırlıkları w_{0k} ($k = 1, \dots, m$) ve sapsmalara uygun birimlerin girdi sinyalleri "1" olarak gösterilmiştir. Ağda sinyallerin yayılımı girdi birimlerinden gizli birimlere sonra ise, gizli birimlerden çıktı birimlerine doğru yönelmiştir. Bu nedenle ağ ileri beslemeli çok katmanlı ağ gibi göze alınmaktadır. Gizli birimin sayısı "1" olduğundan

şekilde gösterilen ağ iki katmanlı bir yapay sinir ağıdır. Gizli birimlerin sayısının birden fazla olduğu yapay sinir ağ mimarileri de vardır (12, 15).



Şekil-12 Tek gizli katmanlı ileri beslemeli çok katmanlı bir yapay sinir ağı

Standart Geriye Yayılım Eğitim Algoritması

Standart geriye yayılım yöntemi üç aşamadan oluşmaktadır: ileri besleme, hatanın hesaplanması ve geriye yayılması, ağırlıkların güncellenmesi.

Geriye Yayılım Algoritma Çeşitleri

Momentum

Momentum geriye yayılımda, ağırlık değişiminin yönü o anki eğitimle bir önceki eğimin kombinasyonu şeklindedir. Bu eğitim azaltma yönteminin değiştirilmiş bir şeklidir ve bazı eğitim verileri, eğitim verilerinin büyük bir çoğunluğundan farklılık gösteriyorsa bu değişim iyi bir avantaj sağlar. Eğer hiç alışılmamış bazı veriler kullanılacaksa bu değişikliği küçük bir öğrenme oranı ile kullanmak iyi olacaktır. Bununla birlikte eğitim verileri benzer olsa da bu değişiklik kullanılarak yaklaşmanın hızı arttırılabilir.

Momentumu kullanmak için bir veya daha önceki eğitim numunelerinin ağırlıkları saklanmalıdır.

Örnek olarak, geri yayılımın momentumlu basit bir biçimi, $t+1$ eğitim adımının yeni ağırlıkları t ve $t-1$ eğitim adımlarındaki ağırlıkları temel alır. Momentumlu geri yayılımın ağırlık güncelleme formülü,

$$w_{jk}(t+1) = w_{jk}(t) + \alpha \delta_k z_j + \mu [w_{jk}(t) - w_{jk}(t-1)] \quad 17.$$

$$\Delta w_{jk}(t+1) = \alpha \delta_k z_j + \mu \Delta w_{jk}(t) \quad 18.$$

$$v_{ij}(t+1) = v_{ij}(t) + \alpha \delta_j x_i + \mu [v_{ij}(t) - v_{ij}(t-1)] \quad 19.$$

$$\Delta v_{ij}(t+1) = \alpha \delta_j x_i + \mu \Delta v_{ij}(t) \quad 20.$$

şeklindedir. Momentum katsayısı değeri olan μ , 0 -1 aralığında sınırlandırılmıştır (12).

Eşlenik Eğitim Algoritması

Birçok eğitim algoritmasında ağırlık güncellenmesinin belirlenmesinde bir öğrenme oranı kullanılmaktadır. Çoğu eşlenik eğitim algoritmalarında ağırlık güncellemesi her bir tekrarda yapılmaktadır. Ağ performans fonksiyonunu minimize eden bu işlem için eşlenik eğitim yönleri boyunca bir arama yapılmaktadır. Birçok arama fonksiyonu vardır. Bazı arama fonksiyonları, belirli eğitim fonksiyonları için çok uygundur.

Çok Katmanlı İleri Beslemeli Ve Geri Yayılımlı Sinir Ağları Ve Onların Öğrenme Algoritmaları

Geri Yayılımın Mantığı

Geri yayılım eğitim yönteminin en genel doğası, geri yayılım tarafından eğitilen çok katmanlı ileri beslemeli bir sinir ağının, çeşitli alanlardaki problemleri çözmek için

kullanabileceğini ifade etmektedir. Örneğin, İngilizce harfleri ve bu harflerden oluşan kelimeleri yüksek sesle okuyabilen *NETtalk* uygulaması, geri yayımlı ağlarla gerçekleştirilmiştir. Bu tür ağları kullanan uygulamalarla, neredeyse her alanda karşılaşılabılır ve bu uygulamalar, doğadaki çeşitli problemler için sinir ağlarını kullanır. Bu problemler, verilen bir girdi grubu için fonksiyon atamayı içerir. Bu girdiler, hedef çıkışların belirli bir grubuna girer. Amaç, çok farklı YSA yapılarında olduğu gibi, aşağıda değinilen şu iki yetenek arasındaki dengeyi elde etmek için YSA'yı, eğitmektir. Bunlardan ilki, girdi numunelerine doğru bir şekilde cevap verebilme yeteneğidir. Burada söz konusu olan girdi numuneleri, eğitim ya da ezberleme için kullanılırlar. İkincisi ise, YSA'nın, sistem içine verilen girdilere karşı uygun çıktıları verebilme yeteneğidir.

Bir YSA'nın geri yayılım yoluyla eğitilmesi üç aşamadan oluşur.

1. Girdi eğitim örneğinin ileri beslenmesi.
2. İlişkili hatanın hesaplanması ve geri yayılımı.
3. Ağırlıkların ayarlanması.

Tek katmanlı bir sinir ağı, fonksiyon atamalarda herhangi bir şekilde sınırlandırılmasına rağmen, bir ya da daha çok saklı katmana sahip çok katmanlı bir YSA, sürekli bir fonksiyon atamayı rasgele bir doğrulukla öğrenebilecektir. Tekbir saklı katmandan daha fazla katmana sahip YSA mimarisi, birçok uygulama için daha yararlı olacaktır.

Çok Katmanlı İleri Beslemeli ve Geri Yayımlı YSA'ların Mimarisi

Y çıkış birimleri ve Z saklı birimleri, şekilde gösterildiği gibi sapmalara sahip olabilir. Genel bir Y_k çıkış birimi üzerindeki sapma w_{ok} ile gösterilir. Tipik bir Z_j saklı birimi üzerindeki sapma da v_{oj} ile gösterilir. Bu sapma terimleri, birimlerden gelen bağlantılar üzerindeki ağırlıklar gibi rol oynar ve söz konusu bu birimlerin çıkışları daima 1'dir.

Çok Katmanlı İleri Beslemeli ve Geri Yayımlı YSA'ların Algoritması

Bir yapay sinir ağının geri yayılım yoluyla eğitilmesi üç aşamadan oluşmaktadır. Bu aşamalar:

- Girdi eğitim örneklerinin ileri beslenmesi,

- İlişkili hatanın geri yayılımı,
- Bağlantı ağırlıkların ayarlanmasıdır.

İleri beslenme boyunca, her bir giriş birimi bir girdi sinyali alır ve bu girdi sinyalini, saklı birimler olan Z_1, \dots, Z_p 'nin her birine yayar. Daha sonra da, her saklı birim, kendi aktivasyonunu hesaplar ve kendi sinyali olan z_j 'yi her bir çıkış birimine gönderir. Y_k çıkış birimlerinin her biri verilen girdi numunesi için ağırlık cevabını biçimlendirmede, kendi aktivasyonu olan y_k 'yi hesaplar.

Eğitim boyunca her bir çıkış birimi, kendi hesaplanmış aktivasyonu olan y_k 'yi, belirli o örnek için, ilişkili hatayı da tespit etmek üzere, hedef çıkış değeri olan t_k ile karşılaştırır ve bu hataya dayalı olan δ_k ($k=1,2,\dots,m$) değerini hesaplar. δ_k değeri, hatayı sınıflara ayırmada Y_k çıkış birimi içinde kullanılır. Hata, önceki katmanda Y_k 'ya bağlanan saklı birimlerden tüm çıkış birimlerine geri döner. O, çıkış katmanı ve saklı katman arasındaki ağırlıkları güncellemek için de kullanılabilir. Benzer şekilde δ_j çarpanı ($j=1,\dots,p$) da, her bir Z_j saklı birimi için ayrı ayrı hesaplanır. Hatayı, geriye yani yatış katmanına doğru yaymak gerekli değildir. Fakat δ_j çarpanı, saklı katman ve giriş katmanı arasındaki ağırlıkları güncellemek için kullanılacaktır.

Çok katmanlı İleri Beslemeli ve Geri Yayılımlı YSA'lar İçin Kullanılan Aktivasyon Fonksiyonları

Çok katmanlı ileri beslemeli ve geri yayımlı sinir ağları için kullanılan aktivasyon fonksiyonlarının önemli şu birkaç özelliğe sahip olmaları gerekir. Aktivasyon fonksiyonu, sürekli, türevlenebilir ve monoton azalmayan bir fonksiyon olmalıdır. Üstelik hesapsal etkinlik için, fonksiyonun türevinin hesaplanmasının oldukça kolay olması beklenir. En yaygın olarak kullanılan aktivasyon fonksiyonları için bağımsız değişkenin belirli bir değerinde, türev değeri fonksiyonun değeri ile ifade edilebilir. Aktivasyon fonksiyonundan, genel olarak, doymuş hale gelmesi de beklenir. Buna örnek olarak, *sonlu maksimum ve sonlu minimum değerlerin asimptotik yaklaşımı kuramı* verilebilir.

Çok Katmanlı İleri Beslemeli ve Geri Yayılımlı YSA'ların Öğrenme Algoritmaları

Uygulamalarda, nöron cevabın girdi değerlerinin sürekli bir fonksiyonu olması gerektiren durumlarda sigmoid ve hiperbolik tanjant gibi aktivasyon fonksiyonları kullanılırken, ikili karar mekanizması gerektiren durumlarda sert geçişli aktivasyon fonksiyonları tercih edilir. Aktivasyon fonksiyonunun parametrik olması, sinaptik ağırlıkların kazandırdığı esnekliğin ötesinde, değişik eşleştirmelerin yapılabilmesini de mümkün kılar.

Lojistik Regresyon Analizi

Sağlık alanındaki araştırmacılar üzerinde çalıştıkları konuda çok etken olması durumunda etkenlerin tek tek bağımlı değişken üzerine etkisi yanında, bunların birlikte etkisini de öğrenmek ya da incelemek istemektedirler. Birlikte etkinin incelenmesinde kullanılan değişik istatistik yöntemler bulunmaktadır. Örneğin, bağımlı değişkenin sürekli, bağımsız değişkenlerin kesikli olması durumunda *varyans analizi*, hepsinin kesikli olması durumunda “*log-linear model*”ler, hepsinin sürekli olması durumunda *regresyon analizi* gibi. Tıp alanındaki araştırmalarda çoğu zaman bağımlı ve bağımsız değişkenlerin tür ve yapıları yukarıda belirtilenlere benzemez, sürekli ve kesikli karışımı bağımsız değişkenlerle karşılaşılır. Üzerinde en çok durulan ve araştırmacı için önemli olan diğer bir konu da etken veya etkenlerle hastalık arasındaki ilişkinin risk yönünden incelenmesidir. Bu tip incelemelerde ağırlıklı olarak LRA kullanılmaktadır (16).

Son yıllarda yoğun bir şekilde kullanılan LRA, gözlemlerin gruplara atanmasında sık kullanılan üç yöntemden (diğerleri kümeleme analizi ve ayırma analizi) birisidir. LRA da grup sayısı bilinmekte, mevcut veriler kullanılarak bir ayırmasama modeli elde edilmekte ve kurulan bu model yardımıyla veri kümesine eklenen yeni gözlemlerin gruplara atanması mümkün olabilmektedir (6).

Doğrusal regresyon analizinde bağımlı değişkenin değeri tahmin edilirken, LRA da bağımlı değişkenin alacağı değerlerden birinin gerçekleşme olasılığı tahmin edilir. Bu olasılık değerinin tahmininde aşağıdaki model kullanılmaktadır.

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad 21.$$

Lojistik regresyon modeliyle tahmin yapılacağı genel olarak kullanılan yöntem en çok olabilirlik metodudur. Genel anlamda en çok olabilirlik metodu, gözlenen veri kümesini elde etmenin olasılığını maksimum yapan bilinmeyen parametrelerin değerlerini tahminlemede kullanılır. Bu metodu uygulayabilmek için en çok olabilirlik fonksiyonunun oluşturulması gerekmekte ve ilgili parametrelerin en çok olabilirlik tahmincileri, fonksiyonu maksimum yapacak değerleri bulacak şekilde seçilmelidir.

LRA’ da gözlenen değerlerle tahmin edilen değerler aşağıdaki ifadeyle karşılaştırılır.

$$D = -2 \ln \left[\frac{\text{Şu andaki modelin olabilirliği}}{\text{Doymuş modelin olabilirliği}} \right] \quad 22.$$

Modelde bulunan herhangi bir bağımsız değişkenin önemliliğine karar vermek için denklemde o bağımsız değişkenin bulunduğu ve bulunmadığı durumlardaki D değerleri, G istatistiği kullanılarak karşılaştırılırlar. G istatistiği p serbestlik derecesiyle ki-kare dağılımı gösterecektir.

$$G = D(\text{Değişkensiz model için}) - D(\text{Değişkenli model için}) \quad 23.$$

$$G = -2 \ln \left[\frac{\text{Değişkensiz modelin olabilirliği}}{\text{Değişkenli modelin olabilirliği}} \right] \quad 24.$$

Katsayıların önemlilikleri test edildikten sonra katsayıların yorumlanması odds oranları kullanılarak yapılmaktadır. LRA' nın kullanım amaçlarından en önemlisi tıp biliminde sıklıkla karşılaşılan bağımlı değişkenin iki ya da daha çok düzey içerdiği, bağımsız değişkenlerin ise hem kesikli hem de sürekli olabildiği durumlarda verilerin ait oldukları gruplara en doğru şekilde atayacak ve hastalıklara ilişkin risk faktörlerini belirleyebilecek modeli kurmaktır. Bunun yanında lojistik regresyon, bağımlı değişkenin tahminini olasılık olarak hesaplayarak olasılık kurallarına uygun sınıflama işlemi yapma imkanı vermektedir (18).

Lojistik Sınıflandırma ve Lojistik Regresyon Modeli

Bağımlı değişkenin 0.1 değerlerine karşılık gelen G_1 ve G_2 grupları, x_1, x_2, \dots, x_p açıklayıcı değişkenlerine dayanılarak sınıflandırılmak istensin. Gruplardaki birey sayısı sıra ile n_1 ve n_2 olduğunda, $N=n_1+n_2$ gözleme dayalı sınıflandırma kuralının oluşumu $f_s(x_1, x_2, \dots, x_p)$ şeklindeki olasılık fonksiyonunun fonksiyonel yapısına ilişkin varsayımlara dayanır. Fonksiyon yapısı için üç tür varsayım söz konusudur.

- Çok değişkenli normal dağılım fonksiyonu
- Lojistik sınıflandırma fonksiyonu
- Dağılımdan bağımsız kernel sınıflandırma fonksiyonu

Lojistik sınıflama fonksiyonu söz konusu olduğunda $X_0=1$ iken $f_s(x_1, x_2, \dots, x_p)$, $G_s(s=1,2)$ grubunun olasılık yoğunluk fonksiyonu olarak tanımlanır. Lojistik varsayım, $\beta'=(\beta_0, \beta_1, \dots, \beta_p)$ için,

$$\frac{f_1(x)}{f_2(x)} = \exp(\beta' X') \text{ ya da } \ln\left(\frac{f_1(x)}{f_2(x)}\right) = \beta' X' \quad 25.$$

Şeklinde tanımlanmaktadır. Bu son eşitlik log-olabilirlik oranı olup x' ler doğrusaldır. Lojistik varsayım bilinmeyen β parametrelerini içermektedir. Her bir gözlem için X koşulu altında gruplardan birine atanma olasılığı olarak tanımlanan sonsal olasılıkları hesaplamak için β tahminleri gerekmektedir. Bunun için lojistik varsayım altında örneklemin olabilirlik fonksiyonu belirlenmelidir. Karışık örneklemede gözlemler (X, G) bileşik dağılımından örneklenmekte yani gözlemler hangi gruptan olduğu bilinmeksizin rasgele seçilmektedir. Buradan G grup üyeliğini gösteren değişken olup iki grup olduğunda G_1 ve G_2 şeklinde gösterilmektedir.

Koşullu örneklemede G' nin x koşulu altında dağılımı incelenmektedir. Biyolojik deneylerin analizinde çok sık kullanılan bu örnekleme türüne ilişkin olabilirlik fonksiyonu diğer örnekleme türlerinin olabilirlik fonksiyonuna temel teşkil etmektedir. Aynı örneklemede ise x' in G koşulu altında dağılımından örnekleme yapılmaktadır. Anderson (1972, 1982) tarafından detaylı olarak incelenen bu örnekleme türünün uygulaması zor olup geçmişe yönelik (retrospective) çalışmalarda uygulanmaktadır (16).

Regresyon problemlerinde anahtar değer, verilen bir bağımsız değişken değerine bağlı olarak bağımlı değişkenin ortalama değerini bulmaktır. Bu değer koşullu ortalama olarak adlandırılır ve $E(Y \mid x)$ ile gösterilir. Burada y' nin bağımlı değişkeni, x' in ise bağımsız değişkeni gösterdiği varsayalım. $E(Y \mid x)$ ifadesi “ x değeri verildiğinde, y' nin beklenen değeri” ni göstermektedir. Doğrusal regresyon analizinde, koşullu ortalamanın, x' in doğrusal bir denklemi olduğu varsayılır.

$$E(Y \mid x) = \beta_0 + \beta_1 x \quad 26.$$

Yukarıdaki bu ifadeden, x' in aralığının $-\infty$ ve $+\infty$ arasından değişmesinden dolayı, $E(Y \mid x)$ 'in mümkün olan her değeri alabileceği görülmektedir. Bağımlı değişken ikili olduğu zaman koşullu ortalama, sıfırla bir arasında değişmek zorundadır.

$[0 \leq E(Y|x) \leq 1]$. x 'deki her birim deęişme sonucunda $E(Y|x)$ 'de oluşan deęişiklik, koşullu ortalama 0'a ya da 1'e yaklaştıkça ilerleyerek az olur.

İki düzey içeren bir bağımlı deęişkenin analizinde kullanılmak üzere önerilen birçok dağılım fonksiyonu bulunmuştur (7). Lojistik dağılım kullanıldığında gösterimi kolaylaştırmak için, x bilindiğinde Y 'nin koşullu ortalamasını göstermek için $\pi(x) = E(Y|x)$ ifadesi kullanılmaktadır. Kullanılacak lojistik regresyon modelinin açık şekli aşağıdaki gibidir.

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad 27.$$

Lojistik regresyon çalışmasına merkez olacak $\pi(x)$ 'in bir transformasyonu yukarıda bahsedildiği gibi lojit transformasyondur. Bu transformasyon $\pi(x)$ cinsinden tanımlanırsa:

$$\begin{aligned} g(x) &= \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] \\ &= \beta_0 + \beta_1 x \end{aligned} \quad 28.$$

Lojit $g(x)$ parametreleri bakımından doğrusal ve x 'in aldığı deęerlere baęlı olarak $-\infty$ ve $+\infty$ arasında deęişmektedir.

Lojistik Regresyon Modelinin Oluşturulması

Lojistik regresyon modeli oluşturulurken, doğrusal regresyonda olduğu gibi maksimum olabilirlik tahmini yöntemi kullanılır. (x_i, y_i) gibi n tane bağımsız gözlem eşinin olduğu varsayıldığında y_i iki düzeyli bağımlı deęişkeni ve x_i 'de i 'inci denek için bağımsız deęişkenin deęerini temsil etsin. Sonuç deęişkeni için 0 ve 1 kodlarının sırasıyla belirli bir olayın varlığını ya da yokluęunu temsil ettiği varsayılınsın. Eşitlik 26'daki lojistik regresyon modelini kestirebilmek için bilinmeyen β_0 ve β_1 parametrelerini tahmin etmek gerekmektedir.

Eğer Y , 0 ve 1 olarak kodlandıysa, $\pi(x)$ ifadesi x verildiğinde Y 'nin 1'e eşit olma koşullu olasılığını vermektedir. $[\pi(x) = P(Y = 1 | x)] \cdot [1 - \pi(x)]$ değeri verilen herhangi bir x için Y 'nin 0'a eşit olma koşullu olasılığını göstermektedir.

$[1 - \pi(x) = P(Y = 0 | x)]$. (x_i, y_i) Çiftinin $y_i = 1$ olduğu zaman olabilirlik fonksiyonuna katkısı $\pi(x_i)$ iken, $y_i = 0$ olduğu zaman olabilirlik fonksiyonuna katkısı $1 - \pi(x_i)$ kadardır. (x_i, y_i) Çiftinin olabilirlik fonksiyonuna katkısını ise aşağıdaki ifade yardımıyla tanımlanmıştır.

$$\zeta(x_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad 29.$$

Gözlemlerin birbirlerinden bağımsız oldukları varsayıldığı için, olabilirlik fonksiyonunu eşitlik 26'daki terimlerin çarpılmasıyla elde edilir.

$$l(\beta) = \prod_{i=1}^n \zeta(x_i) \quad 30.$$

En çok olabilirliğin temel ilkesinde β tahmininin eşitlik 28'deki ifadeyi maksimum yaptığı vurgulanır, öte yandan log olabilirlik fonksiyonu ise:

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\} \quad 31.$$

$L(\beta)$ 'yi maksimum yapan β değerini bulabilmek için, $L(\beta)$ 'nin β_0 ve β_1 'e göre türevini aldıktan sonra, elde edilen ifadeler 0'a eşitlenmelidir. Sonuçta elde edilen en çok olabilirlik eşitlikleri aşağıdaki gibidir.

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad 32.$$

$$\sum_{i=1}^n x_i [y_i - \pi(x_i)] = 0 \quad 33.$$

Eşitlik 32 ve 33'den elde edilen β 'nin değeri, en çok olabilirlik tahmincisi olarak adlandırılır ve β olarak gösterilir. Örnek olarak, $\pi(x_i)$ 'nin en çok olabilirlik tahmini

$\pi(x_i)$ ile gösterilir. Bu değer, x' in x_i gibi bir değere eşit olduğu verildiği zaman, Y' nin 1'e eşit olma koşullu olasılığının tahminini verir.

$$\sum_{i=1}^n y_i = \sum_{i=1}^n \hat{\pi}(x_i) \quad 34.$$

Lojistik Regresyon Modelindeki Katsayıların Önem Testleri

Modeldeki bir değişkenin katsayısının önem testi için bir yaklaşımı. “Soru altındaki değişkeni kapsayan model, sonuç değişkeni hakkında o değişkeni kapsamayan modelden daha çok bilgi içeriyor mu?” sorusu ile ilgilidir. Bu soru bağımlı değişkenin gözlenen değerlerini, her iki modelden elde edilen tahmini değerlerle karşılaştırılarak cevaplanır. Eğer değişkenli modelin tahmin edilen değerleri değişkeni kapsamayan modelden daha iyi ya da herhangi bir yönden daha açıksa o zaman önemliliği araştırılan değişkenin önemli olduğu sonucuna varılır. Burada dikkat edilmesi gereken önemli bir noktada tahmin edilen değerlerin, gözlenen değerleri ne kadar doğru yansıttığı sorusu değildir (bu uyum iyiliği olarak adlandırılır).

Doğrusal regresyonda eğim katsayısının önemine karar verme işlemi varyans analizi tablosunun açıklanmasıyla mümkün olmaktadır. Bu tablonun bölümleri gözlemlerin ortalamalarından ayrılışlarının kareler toplamını ikiye böler:

- Regresyon çizgisinden gözlemlerin ayrılış kareleri toplamı *RAKT*, (ya da regresyondan ayrılış kareler toplamı)
- Bağımlı değişkenin ortalamasından, regresyon modelinden elde edilen tahmin edilen değerlerin ayrılış kareleri toplamı *RKT*, (ya da regresyon kareler toplamı).

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

35.

$$RKT + RAKT$$

Doğrusal regresyonda gözlenen değerlerin, tahmin edilen değerlere göre karşılaştırılması her ikisi arasındaki uzaklığın karelerine dayanır. Eğer i inci kişi için, y_i

gözlenen değeri ve \hat{y}_i ise tahmin edilen değeri gösterirse, bu karşılaştırmayı değerlendirmek için kullanılan istatistik aşağıdaki gibidir.

$$RKT = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad 36.$$

Modelde bağımsız değişken olmadığı zaman tek parametre vardır oda β_0 'dır ve $\beta_0 = \bar{y}$ cevap değişkeninin ortalamasına eşittir. Bu durumda $\hat{y}_i = \bar{y}$ ve *RAKT* toplam varyansa eşittir. Bağımsız değişken modele dahil edildiğinde ise, *RAKT* da meydana gelen herhangi bir azalma, bağımsız değişken için eğim katsayısının sıfırdan farklı olmasından kaynaklanmaktadır.

$$RKT = \left[\sum_{i=1}^n (y_i - \bar{y})^2 \right] - \left[\sum_{i=1}^n (y_i - \hat{y}_i)^2 \right] \quad 37.$$

Lojistik regresyonda katsayıların önem testi için yol gösterici ilke bahsedildiği gibi sorgulama altındaki değişkeni kapsayan ve kapsamayan modellerden elde edilen tahmin değerlerini, bağımlı değişkenin gözlenen değerleriyle karşılaştırmaktır. Gözlenen ve tahmin edilen değerlerin karşılaştırılma işlemi log-olabilirlik fonksiyonu ile yapılır. Bu karşılaştırmayı daha iyi anlayabilmenin bir yolu, doymuş modelden elde edilen tahmin edilen değerleri, bağımlı değişkenin gözlenen değerleri olarak kabul etmektir. Doymuş model veri noktası kadar parametre içeren modeldir.

Çoklu Lojistik Regresyon Modeli

Birden çok bağımsız değişkenin yer aldığı Lojistik modellere çok değişkenli lojistik regresyon adı verilir. Yapısal olarak bu modelin diğer çok değişkenli regresyon modellerinden farkı olmayıp regresyon katsayılarının yorumlanması farklıdır. Çoklu lojistik regresyon modelinde genel eğilim modeldeki katsayıların tahmini ve onların önem testi şeklinde olmaktadır.

$x' = (x_1, x_2, \dots, x_p)$ vektörü ile gösterilen, p tane bağımsız değişken toplandığı varsayalım. Şu an için bu değişkenlerin her birinin en azından aralıklı ölçekle (sürekli) ölçüldüğü varsayalım. Bağımlı değişkeninin mevcut olduğu ($Y=1$) zaman ki koşullu

olasılık $P(Y=1|x) = \pi(x)$ 'e eşittir. Çoklu lojistik regresyon modelinin lojiti aşağıdaki eşitlik ile verilmiştir,

$$g(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \dots \quad 38.$$

bu durumda

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad 39.$$

Eğer bazı bağımsız değişkenler kesikli, nominal ölçekli ise, o zaman bu değişkenleri sürekli değişkenlermiş gibi modele dahil etmek yanlış olacaktır. Çeşitli düzeyleri göstermek için kullanılan sayılar sadece tanımlayıcıdır ve bunların herhangi bir sayısal değeri yoktur. Bağımsız değişkenler sayısal olarak sınıflandırıldığı zaman çeşitli dizayn değişkenlerinin (kukla değişken) kategorik olan bu değişkenleri temsil etmesi için kullanılması gerekmektedir (17).

İstatistik paket programlarının bir çoğu bu tür dizayn değişkenlerini kullanmaktadır ve bazı programlarda birkaç farklı metod seçeneği vardır. Örneğin SPSS paket programında prosedürlerde referans grup ilk ya da sonraki olarak belirlenebilmektedir.

1. **Tekrarlı (repeated) Yöntem:** Herbir grubun riskini, kendinden öncekine göre hesaplanmasına olanak sağlar.
2. **Fark (difference) Yöntemi:** Herbir grubun riskini, kendinden öncekilerin ortalama riskine göre hesaplanmasını sağlar.
3. **Helmert Yöntemi:** Herbir grubun riskini, kendinden sonrakilerin ortalama riskine göre hesaplanmasını sağlar.
4. **Sapma (deviance) Yöntemi:** Herbir grubun riskini, grupların toplam riskine göre hesaplanmasını sağlar.
5. **Basit (simple) Yöntem:** Herbir grubun riskini, referans (temel) gruba göre hesaplanmasını sağlar.

Sağlık bilimlerinde yukardaki yöntemlerden, en çok kullanılanı **basit yöntem**'dir. Bu yöntemin tercih edilmesinin nedeni epidemiyolojik ve klinik amaçlara uygun olmasıdır (16).

Genel olarak, eğer nominal bir değişken k kategoriye sahipse, o zaman $k-1$ dizayn değişkenine ihtiyaç vardır. j ' ninci bağımsız değişken (x_j) k_j kategoriye sahip olsun. $k_j - 1$ dizayn değişkeni D_{ju} olarak katsayıları da β_{ju} , $u = 1, 2, \dots, k_j - 1$ olarak belirtilmiş olsun. Sonuç olarak J ninci değişkeni kesikli olan p değişkenli model için lojit aşağıdaki gibidir.

$$g(x) = \beta_0 + \beta_1 x_1 + \dots + \sum_{u=1}^{k_j-1} \beta_{ju} D_{ju} + \beta_p x_p \dots \quad 40.$$

Çoklu Lojistik Regresyon Modelinin Kurulması

Birbirinden bağımsız n tane (x_i, y_i) , $i = 1, 2, \dots, n$ gözlem eşinin olduğu varsayalım. Tek değişkenli modelde olduğu gibi modelin kurulması için tahmin vektörünün $\beta' = (\beta_0, \beta_1, \dots, \beta_p)$ elde edilmesi gerekmektedir. Çok değişkenli durumda, tek değişkenli durumda olduğu gibi tahmin metodu en çok olabilirlik olacaktır. Log olabilirlik fonksiyonunun $p-1$ katsayısına göre türevi alınarak, $p+1$ tane olabilirlik denklemi elde edilebilmektedir.

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad 41.$$

$$\sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0, j = 1, 2, \dots, p \dots \quad 42.$$

Modelin Önemlilik Testinin Yapılması

Çok değişkenli lojistik regresyon modelinin oluşturulmasından sonraki aşama modelde bulunan değişkenlerin önemliliklerinin araştırılmasıdır. Olabilirlik oran testi, tek değişkenli durumda olduğu gibi, modeldeki bağımsız değişkenler için p katsayısının tümel önemliliğini vermektedir. Olabilirlik oran testi, ilgili eşitliklerde verildiği gibi G istatistiğine bağlıdır. "Modeldeki birlikte değişenler için p tane eğim katsayısının sıfıra eşit olması" hipotezi altında, G istatistiği p serbestlik derecesiyle ki-kare dağılımına sahip olacaktır. H_0 hipotezinin red edilmesi durumunda en az bir tane katsayının sıfırdan farklı olduğu sonucuna varılır.

Katsayılarının hepsinin ya da bir kısmının sıfırdan farklı olduğu sonucuna varmadan önce, değişkenler tek tek Wald test istatistiği ile test edilebilir

$$|W_j = \hat{\beta}_j / \hat{SE}(\hat{\beta}_j)|$$

Dikkat edilmesi gereken bir nokta da, en iyi uyuma sahip modeli en az parametre ile belirlemektir. Bundan sonraki ilk adım, sadece önemli olduğu düşünülen değişkenleri modele dahil ederek yeni bir model oluşturmak ve bu modeli tüm değişkenleri kapsayan tam modelle karşılaştırmaktır. Bu iki model olabilirlik oran testi kullanılarak karşılaştırılır. “Modelden çıkarılan değişkenlerin katsayıları sıfıra eşittir “ hipotezi altında G istatistiği ($v_2 - v_1$) serbestlik derecesiyle ki-kare dağılımı göstermektedir,

v_2 : (Tam modeldeki değişken sayısı +1),

v_1 : (Azaltılmış modeldeki değişken sayısı +1) olarak verilmiştir.

$$G = -2 [(\text{Azaltılmış model için log-olabilirlik}) - (\text{Tam model için log-olabilirlik})]$$

($v_2 - v_1$) serbestlik derecesinde yanılma olasılığı (p) bulunabilir. Eğer yanılma olasılığı (p) 0.05’den büyük olursa, değişken sayısı azaltılmış modelin tam model kadar iyi olduğuna karar verilmektedir. Kategorik olarak ölçklendirilmiş bir bağımsız değişken modelden çıkarıldığı (ya da modele dahil edildiği) zaman, onun bütün dizayn değişkenleri de modelden çıkarılmalıdır (ya da dahil dilmelidir). Eğer kategorik bir değişkenin k düzeyi varsa, bu değişkenin modelden çıkarılmasının olabilirlik oran testi için serbestlik derecesine katkısı $k-1$ olacaktır. Çoklu serbestlik derecesinden dolayı katsayıların önemliliğini araştırmak için Wald istatistiği kullanılırken dikkatli olunması gerekmektedir.

Wald testinin çok değişkenli karşılığı aşağıdaki vektör –matris ifadeleri yardımıyla elde edilebilmektedir.

$$\begin{aligned} W &= \hat{\beta}' [\hat{\Sigma}(\hat{\beta})]^{-1} \hat{\beta} \\ &= \hat{\beta}' (X' V X) \hat{\beta} \end{aligned} \quad 43.$$

$p+1$ katsayının her birinin sıfıra eşit olması hipotezi altında W istatistiği $p+1$ serbestlik derecesinde ki-kare dağılımına sahip olacaktır. p eğim katsayısı için testler $\hat{\beta}$ vektöründen $\hat{\beta}_0$ ’ı ve $\hat{\beta}_0$ ’le ilgili satır (ilk satır) ve sütunu (ilk sütun), $(X' V X)$ matrisinden eleyerek elde edilmektedir. Bu testin gerçekleştirilmesi için vektör-matris işlemlerini

gerçekleştirebilecek ve $\hat{\beta}$ ' i elde edebilecek yeterlilik gerektiğinden dolayı yine olabilirlik oran testinden daha fazla

Lojistik Regresyon Modelinde Katsayıların Yorumlanması

Karar vermemiz gereken ilk adım, “bağımlı değişkenin hangi fonksiyonu bağımsız değişkenler ile doğrusal bir fonksiyon oluşturmaktadır?” sorusudur. Bu fonksiyona link fonksiyonu adı verilir.

Doğrusal regresyon modelinde link fonksiyonu I (identity-Birim) matristir. Çünkü bağımlı değişken parametreleri ile doğrusaldır. Lojistik regresyon modelinde ise link fonksiyonu lojit transformasyondur (16).

$$g(x) = \ln \left\{ \frac{\pi(x)}{[1 - \pi(x)]} \right\} = \beta_0 + \beta_1 x \quad 44.$$

Modelde Sadece İkili Bağımsız Değişkenin Olduğu Durum

Lojistik regresyon katsayılarının yorumuna bağımsız değişkenin ikili olduğu zamanki durum ile başlanacaktır. Bu durum hem en basit durum olup, hem de diğer bütün durumlar için detaylı bir temel teşkil edecektir. x 'in 0 ve 1 olarak kodlandığı varsayalım. Bu model altında $\pi(x)$ ve $1 - \pi(x)$ 'in ikişer değerleri vardır. $\pi(x)$ için bu değerler $\pi(0)$ ve $\pi(1)$ iken, $1 - \pi(x)$ için bu değerler $1 - \pi(0)$ ve $1 - \pi(1)$ 'dir. Bu değerler 2x2 tabloda aşağıdaki gibi gösterilmiştir.

Tablo-1 Bağımsız değişkenin ikili olduğu zaman lojistik regresyon modelinin değerleri

| | x=1 | x= 0 |
|---------------|--|--|
| y=1 | $\pi(1) = \frac{e^{\beta_0 + \beta_1}}{1 + e^{\beta_0 + \beta_1}}$ | $\pi(0) = \frac{e^{\beta_0}}{1 + e^{\beta_0}}$ |
| y=0 | $1 - \pi(1) = \frac{1}{1 + e^{\beta_0 + \beta_1}}$ | $1 - \pi(0) = \frac{1}{1 + e^{\beta_0}}$ |
| Toplam | 1.0 | 1.0 |

Modelde İki Den Fazla Düzeyli Bağımsız Değişkenin Olduğu Durum

Bağımsız değişkenin ikiden fazla düzey içerdiği durumlarda olabilir ($k > 2$). Bu tür nominal ölçekli bağımsız değişkenleri sürekli değişkenlermiş gibi modele dahil etmek doğru değildir. Bunun için ikiden çok kategorisi olan bağımsız bir değişkenin, dizayn değişkenleri kullanılarak modele dahil edilmesi gerekmektedir.

Modelde Sürekli Bir Bağımsız Değişkenin Olduğu Durum

Sürekli bir değişken için katsayının yorumlanması amacıyla geliştirilecek metod için lojitin değişkenle doğrusal olduğu varsayılacaktır. Lojitin sürekli değişkenle (x) doğrusal olması varsayımı altında lojit için eşitlik $g(x) = \beta_0 + \beta_1 x$ 'dir. Eğim katsayısı (β_1) x 'deki "1" birimlik artışın log odds değerinde meydana getireceği değişimi verir.

x 'deki "c" birimlik bir değişim için log odds değeri lojit farktan elde edilmektedir. $g(x + c) - g(x) = c\beta_1$. Karşılık gelen odds oranı lojit farkın üssü alınarak elde edilir. $\psi(c) = \psi(x + c, x) = \exp(c\beta_1)$ $\psi(c)$ 'nin tahmini için $\%100(1 - \alpha)$ güven aralığının uç noktaları

$$\exp\left[c\hat{\beta}_1 \pm z_{1-\alpha/2}c\hat{SE}(\hat{\beta}_1)\right]$$

45.

olarak verilmiştir.

Modelin Çok Değişkene Sahip Olması

Çok değişkenli durumda modelde yalnızca iki bağımsız değişken olduğu (biri ikili, diğeri sürekli) varsayalım. Öncelikli olarak ilgi ikili bağımsız değişkenin etkisi üzerine olacaktır. Risk faktörüne maruz olma “var ya da yok” olarak kodlandığında, sürekli bir değişken için ayarlama yapılması epidemiyolojik araştırmalarda sıkça karşılaşılan bir durumdur. Doğrusal regresyondaki benzer bir durum kovaryans analizi olarak adlandırılmaktadır.

Etkileşim ve Etki Karışımı

Etkileşim, kontrol değişkeninin farklı düzeylerinde bağımsız değişkenle bağımlı değişken arasındaki ilişkinin farklı olmasıdır. Bu durumda kontrol değişkeni, bağımsız değişkenin bağımlı değişken üzerindeki etkisinde değişikliğe neden olur. Bu durumu yaratan kontrol değişkenine etki değiştirici adı verilir. Etkileşimin anlamlı olması halinde bağımsız değişkene ilişkin risk oranları kontrol değişkeninin her bir düzeyi için ayrı değerlendirilip yorumlanmalıdır. Etki değiştirici bir kontrol değişkeninin etkisi, tabakalı çalışma yapılarak kontrol edilebilir.

Etki karışımı, kavram olarak etkileşimle karışmaktadır, ancak etkileşim ve etki karışımı aynı değildir. Etki karışımının varlığını anlamak etkileşim kadar kolay değildir. Etkileşimde kontrol değişkeninin her bir düzeyinde bağımsız-bağımlı değişken ilişkisi farklı iken etki karışımında bu özellik yoktur. Kontrol değişkeninin etkisi bağımsız-bağımlı değişken ilişkisi içinde gizlidir. Bu gizlilik kontrol değişkeninin hem bağımlı hem de bağımsız değişkenle ilişkisinden kaynaklanmaktadır. Bu nedenle bağımsız değişkenle, bağımlı değişken ilişkisinin etkileşimde olduğu gibi kontrol değişkeninin her bir düzeyine göre incelenmesi, kontrol değişkeninin etkisini gidermez. Etki karışımını gidermek için kontrol değişkeninin etkisini sabit tutmak gerekir (16).

GEREÇ VE YÖNTEM

LRA ve YSA tekniklerinin sınıflama etkinliklerinin karşılaştırılması amacıyla, gereksinim duyulan veri seti Ercan ve arkadaşları (1) tarafından yapılan çalışmanın veri setinden çalışmamıza uygun 140 klinik hastası alınmıştır. Çalışmamızın konusu kapsamında yatan hasta memnuniyeti örnek sınıflama problemine temel oluşturacak şekilde seçilmiş ve yine bu doğrultuda konunun uzmanı tarafından ilgili ölçekten 7 tane değişken, analizi yapılacak olan veri setine dahil edilmek üzere seçilmiştir.

Bağımlı değişken, yatan hasta memnuniyetidir, yatan hastanın memnuniyetini etkilediği düşünülen etkenler ise bağımsız değişkenler olarak alınmıştır ve veri setini oluşturan tüm değişkenler aşağıdaki tabloda verilmiştir.

| | | | |
|----------------------|---|--|--|
| Y | : | Yatan hasta memnuniyeti | (memnun değil, memnun) |
| X₁ | : | Gelir düzeyi | (≤ 600 YTL , > 600 YTL) |
| X₂ | : | Eğitim düzeyi | (ilköğretim mezunu ise, lise mezunu ise, üniversite mezunu ise) |
| X₃ | : | Yaş | (≤ 30 , > 30) |
| X₄ | : | Son bir yıl içinde hekime gitme sayısı | (≤ 4 , > 4) |
| X₅ | : | Odada beraber kalınan hasta sayısı | (≤ 2 kişi , > 2) |
| X₆ | : | Hastanın en son yatışında hastanede kaldığı gün sayısı | (≤ 3 gün , > 3 gün) |
| X₇ | : | En son yatarak tedavi olduğu zaman süreci | (< 3 ay , 3 ay -12 ay , > 12 ay) |

Çözüm süresince, LRA için SPSS 13.0 YSA analizi için ise STATISTICA 6.0 istatistiksel analiz paket programları kullanılmıştır.

BULGULAR

Uygulamaya dâhil edilen 140 klinik hastası 2002 yılında Ercan ve arkadaşları(1) tarafından yapılan çalışma temel alınarak “memnun” ve “memnun değil” olarak sınıflandırılmıştır. Bu sınıflandırma için LRA ve YSA teknikleri kullanılmış ve sonuçta iki yöntemden elde edilen doğru sınıflandırma oranları karşılaştırılmıştır. Yatan hasta memnuniyetiyle ilişkili olan risk faktörleri, çok değişkenli LRA ile belirlenmiştir. Çalışmaya alınan 140 deneğe ait tanımlayıcı bilgiler aşağıdaki tablolarda verilmiştir.

Tablo-2 Deneklerin hasta memnuniyetine göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|--------------|------|-----------|
| Memnun | 105 | 75 |
| Memnun Değil | 35 | 25 |

Tablo-3 Deneklerin gelir durumlarına göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|---------------------|------|-----------|
| 600 Milyondan Az | 100 | 71.4 |
| 600 Milyondan Fazla | 40 | 28.6 |

Tablo-4 Deneklerin eğitim durumlarına göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|-------------------|------|-----------|
| İlköğretim mezunu | 73 | 52.1 |
| Lise mezunu | 40 | 28.6 |
| Üniversite mezunu | 27 | 19.3 |

Tablo-5 Deneklerin yaş gruplarına göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|-------------------------|------|-----------|
| 30 yaşına eşit ve küçük | 86 | 61.4 |
| 30 yaşından büyük | 54 | 38.6 |

Tablo–6 Deneklerin son bir yıl içinde hekime gitme sayılarına göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|---------------|------|-----------|
| 4 ve 4'ten az | 67 | 47,9 |
| 4'ten fazla | 73 | 52.1 |

Tablo–7 Deneklerin son odayı paylaştıkları hasta sayılarına göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|---------------|------|-----------|
| 2 ve 2 den az | 73 | 52.1 |
| 2 den fazla | 67 | 47.9 |

Tablo–8 Deneklerin en son hastanede yattıkları gün sayısına göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|---------------|------|-----------|
| 3 ve 3'ten az | 94 | 67.1 |
| 3'ten fazla | 46 | 32.9 |

Tablo–9 Deneklerin en son hastanede tedavi olma durumlarına göre dağılımları

| Değişken | Sayı | Yüzde (%) |
|---------------------|------|-----------|
| 3 aydan az | 43 | 30.7 |
| 3 ay ve 12 ay arası | 49 | 35 |
| 12 aydan fazla | 48 | 34.3 |

LRA ile Sınıflandırma

Yatan hasta memnuniyeti olarak belirlenen Y bağımlı değişkeni, memnuniyet durumunda 1, memnun olmama durumunda ise 0 olarak kodlanmıştır. LRA ile test edilecek olan hipotezler ise aşağıdaki şekilde kurulmuştur.

H_0 : Gelir düzeyinin, eğitim düzeyinin, yaşının, son bir yıl içinde hekime gitme sayısının, odada beraber kaldığı hasta sayısının, en son yatışında hastanede kaldığı gün sayısının, en son yatarak tedavi olduğu zaman sürecinin, hasta memnuniyeti üzerinde bir etkisi yoktur.

H_1 : En az bir değişkenin hasta memnuniyetine etkisi vardır.

LRA da, bağımlı değişken olarak alınan yatan hasta memnuniyeti değişkenini etkileyen bağımsız değişkenlerin belirlenmesi aşamasında Wald testi kullanılmıştır. \hat{b} parametreleri ile bu parametrelere ilişkin Wald istatistikleri, serbestlik dereceleri, anlam düzeyleri, $\text{Exp}(\hat{b})$ değerleri Tablo-9’da verilmiştir.

Tablo-9 LRA sonucunda elde edilen değerler

| Değişken | \hat{b} | Standart Hata | Wald | s.d. | p | Exp(\hat{b}) |
|---|-----------|----------------------|-------------|-------------|----------|----------------------------------|
| Sabit Değer | -1.410 | 0.793 | 3.161 | 1 | 0.075 | 0.244 |
| X₁, Gelir Düzeyi <i>>600 YTL</i> | 0.426 | 0.621 | 0.469 | 1 | 0.493 | 1.531 |
| X₂, Eğitim Durumu | | | 5,357 | 2 | 0.039 | |
| X₂₁, Eğitim Durumu <i>Lise Mezunu</i> | 1.449 | 0.723 | 4.012 | 1 | 0.045 | 1.435 |
| X₂₂, Eğitim Durumu <i>Üniversite Mezunu</i> | 0.361 | 0.636 | 0.323 | 1 | 0.040 | 3.258 |
| X₃₁, Yaş <i>>30</i> | 0.172 | 0.475 | 0.132 | 1 | 0.717 | 1.188 |
| X₄, Hekime Gitme Sayısı (Son 1 yıl içinde) <i>>4</i> | 0.448 | 0.464 | 0.931 | 1 | 0.335 | 1.565 |
| X₅₁, Odada beraber kalınan hasta sayısı <i>>2 kişi</i> | 0.610 | 0.471 | 1.679 | 1 | 0.022 | 2.840 |
| X₆₁, Hastanın en son yatışında hastanede kaldığı gün sayısı <i>>3 gün</i> | 1.006 | 0.452 | 4.966 | 1 | 0.026 | 2.736 |
| X₇, En son yatarak tedavi olduğu zaman süreci | | | 0.924 | 2 | 0.169 | |
| X₇₁, En son yatarak tedavi olduğu | 0.130 | 0.532 | 0.060 | 1 | 0.1807 | 1.139 |

| | | | | | | |
|---|-------|-------|-------|---|-------|-------|
| zaman süreci 3 ay – 12 ay | | | | | | |
| X₇₂, En son yatarak tedavi olduğu zaman süreci > 12 ay | 0.579 | 0.632 | 0.839 | 1 | 0.360 | 1.784 |

Hosmer ve Lemeshow lojistik regresyonda Wald testi için anlam düzeyi olarak 0.15 veya 0.25 kullanmayı önermişlerdir (19, 20). Çalışmamızda anlam düzeyi olarak $p = 0.15$ alınmıştır. Buna göre bağımlı değişkeni etkileyen değişkenler, önem düzeyleri 0.15'ten küçük olan X_2 (eğitim durumu) [$p = 0.069$], X_5 (odada beraber kalınan hasta sayısı) [$p = 0.022$] ve X_6 'dır (Hastanın en son yatışında hastanede kaldığı gün sayısı) [$p = 0.026$].

Elde edilen OR değerlerine göre, üniversite mezunu olanlar ilköğretim mezunlarına göre 3.258 kat, lise mezunları ise yine ilköğretim mezunlarına göre 1.435 kat daha fazla memnun olmama riskine sahiptirler. Kaldığı odayı 2 den fazla kişiyle paylaşanlar 2 den az kişiyle paylaşanlara göre 2.840 kat daha fazla memnun olmama riskine sahiptirler. Hastanın en son yatışında hastanede kaldığı gün sayısına göre 3 günden fazla kalanlar, 3 günden az kalanlara göre 2.736 kat daha fazla memnun olmama riskine sahiptirler.

LRA sonucunda lojistik regresyon modeli

$$\hat{Y} = \frac{1}{1 + e^{-[-1,410 + 0,426X_1 + 1,449X_{21} + 0,361X_{22} + 0,172X_{31} + 0,448X_4 + 0,610X_{51} + 1,006X_{61} + 0,130X_{71} + 0,579X_{72}]}} \quad 46.$$

olarak elde edilmiştir.

Elde edilen modelin geçerliliği Hosmer Lemeshow testi ile sınanmıştır.

H_0 : Tahmin denklemi anlamlıdır.

H_1 : Tahmin denklemi anlamlı değildir.

Elde edilen tahmin modeli için Hosmer Lemeshow ki-kare değeri 7.536 olarak hesaplanmıştır. $p = 0.462 > \alpha = 0.05$ olarak elde edilmiş ve modelin uygun olduğuna dair

H_0 hipotezi kabul edilmiştir. Buna göre lojistik regresyon sonucunda elde edilen bağımsız değişkenlerden en az birinin katsayısı sıfırdan farklı çıkmıştır. Bulunan tahmin denklemi anlamlıdır.

46 numaralı eşitlik ile sınıflandırma yapılmaktadır. Bu işlem için bağımlı değişken tahmin değerleri hesaplanarak, 0.5'ten küçük olanlara "0" değeri ve 0.5'ten büyük olanlara "1" değeri atanmıştır. Yapılan bu atama değerleri gözlem değerleri ile karşılaştırılarak doğru sınıflandırma oranları hesaplanmıştır. Elde edilen sonuçlar Tablo-10'de verilmiştir:

Tablo-10 LRA sınıflandırma tablosu

| | | | Tahmin Değerleri | | | |
|-----------------|--------|---------|------------------|-------|--------|-----------------|
| | | | Memnun | | | |
| | | | Ol mama | Ol ma | Toplam | Doğruluk |
| Gerçek Değerler | Memnun | Ol mama | 12 | 23 | 35 | %34.3 |
| | | Ol ma | 3 | 102 | 105 | %97.1 |
| Toplam | | | 15 | 125 | 140 | %81.4 |

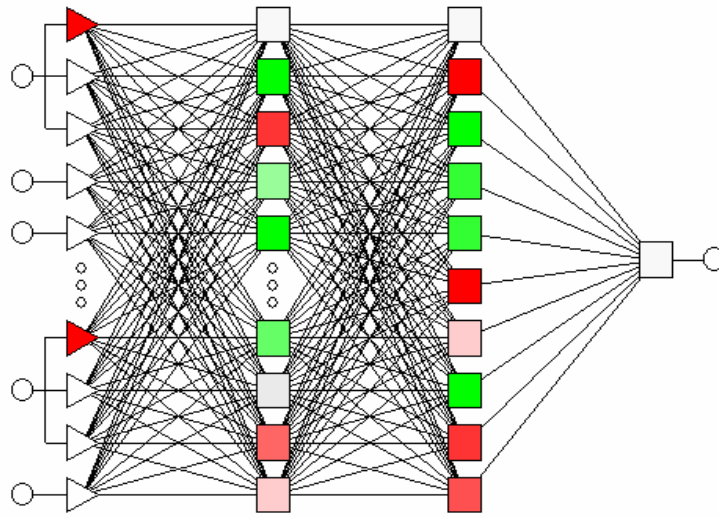
Tablo-10' da görüldüğü gibi, uygulamadaki 140 hastadan gerçekte 35 tanesi memnun değildir. Geriye kalan 105 hasta ise memnundur. Memnun olmayan 35 hastanın 12 tanesi LRA ile yapılan sınıflandırma işleminde doğru, 23 tanesi ise hatalı olmak üzere % 34.3'lük bir doğruluk yüzdesi ile sınıflandırılmıştır. Memnun olan 105 hastanın ise 102 tanesi LRA ile yapılan sınıflandırma işleminde doğru, 3 tanesi ise hatalı olmak üzere % 97,1'lik bir doğruluk yüzdesi ile sınıflandırılmıştır. LRA ile yapılan sınıflandırma işleminde genel doğruluk değeri ise 140 hastanın 114 tanesi doğru sınıflandırılarak % 81.4 olarak hesaplanmıştır.

Yapay Sinir Ağı Modeli ile Sınıflandırma

Uygulamada kullanılan veri setinin sınıflandırılması için girdi katmanında 7 birim, çıktı katmanında 1 birim olan ve iki gizli katmana sahip, ilk gizli katmanında 11 ikinci

gizli katmanında 10 birim olan üç katmanlı bir yapay sinir ağı modeli kurulmuştur. Veri seti, %30'u eğitim, %30'u geçerlilik ve %40'ı test seti olmak üzere üçe ayrılmıştır. En yüksek sınıflandırma oranını elde edebilmek için gizli katmanlar ve sahip oldukları birim sayıları simülasyon yapılarak 1000 tane yapay sinir ağı modeli denenmiştir. Memnuniyet değişkeni iki kategoriye (0.1) sahip olduğundan gizli katman ve çıktı katmanında aktivasyon fonksiyonu olarak sigmoid lojistik fonksiyonu kullanılmıştır. Kurulan model şekil-13'de gösterilmiştir:

Train Perf. = 0,917808 , Select Perf. = 0,742857 , Test Perf. = 0,771429



Şekil-13 Yapay sinir ağı analizi sonucunda elde edilen model

Şekil 13'de gösterilen yapay sinir ağında girdi katmanında 7 adet, gizli katmanlarından birincisinde 11 ikincisinde 10 adet ve çıktı katmanında 1 adet nöron bulunmaktadır. Gizli ve çıktı katmanların aktivasyon fonksiyonları olarak lojistik fonksiyon kullanılmaktadır.

Şekil 13'de gösterilen yapay sinir ağı, danışmanlı öğrenme yöntemlerinden olan geriye yayılım algoritmasıyla eğitilmiştir. Eğitim esnasında momentumlu geriye yayılım eğitim algoritması yöntemi kullanılmıştır. Kurulan model yardımıyla eğitim, geçerlilik ve test setinden elde edilen doğru sınıflandırma oranları tablo-11' de verilmiştir.

Tablo–11 Kurulan model yardımıyla eğitim, geçerlilik ve test setinden elde edilen doğru sınıflandırma oranları

| | Eğitim Seti | Geçerlilik Seti | Test Seti |
|-------------------------|-------------|-----------------|-----------|
| Sınıflandırma Oranı (%) | 0.917808 | 0.742857 | 0.771429 |

Ağın eğitimi için 1000 iterasyon gerçekleştirilmiştir. YSA analizi sonucunda doğru sınıflandırma oranı % 85.7 olarak bulunmuştur. Sınıflandırma tablosu tablo-12’de gösterilmiştir.

Tablo–12 YSA sonucu elde edilen doğru sınıflandırma oranı tablosu

| | | | Tahmin Değerleri | | | |
|-----------------|--------|--------|------------------|------|--------|-----------------|
| | | | Memnun | | | |
| | | | Olmama | Olma | Toplam | Doğruluk |
| Gerçek Değerler | Memnun | Olmama | 26 | 9 | 35 | %74.3 |
| | | Olma | 12 | 93 | 105 | %88.6 |
| Toplam | | | 38 | 102 | 140 | %85 |

Tablo–12 de görüldüğü gibi, uygulamadaki 140 hastadan gerçekte 35 tanesi memnun değildir. Geriye kalan 105 hasta ise memnundur. Memnun olmayan 35 hastanın 26 tanesi YSA ile yapılan sınıflandırma işleminde doğru, 9 tanesi ise hatalı olmak üzere % 74.3’lük bir doğruluk yüzdesi ile sınıflandırılmıştır. Memnun olan 105 hastanın ise 93 tanesi YSA ile yapılan sınıflandırma işleminde doğru, 12 tanesi ise hatalı olmak üzere % 88.6’lık bir doğruluk yüzdesi ile sınıflandırılmıştır. YSA ile yapılan sınıflandırma işleminde genel doğruluk değeri ise 140 hastanın 119 tanesi doğru sınıflandırılarak % 85 olarak hesaplanmıştır. Uygulama sonucunda elde edilen değerler tablo–13 de verilmiştir.

Tablo-13 LRA ve YSA analizi ile elde edilen doğru sınıflandırma oranları

| Model | Memnun | | Genel % |
|-------|------------|------------|-------------|
| | Olmama | Olma | |
| | Doğruluk % | Doğruluk % | |
| LRA | 34.3 | 97.1 | 81.4 |
| YSA | 74.3 | 88.6 | 85 |

Tablo-13' den de görülebileceği gibi verilerin doğru sınıflandırma oranları LRA için % 81.4 ve YSA için % 85' tir.

TARTIŞMA VE SONUÇ

Çok deęişkenli istatistiksel verilerin sınıflandırılması, bu verilere uygulanabilecek çeşitli istatistiksel yöntemler için gerekli ve yararlı bilgiler verecektir (16). Birimleri sınıflandırma yöntemlerinden ikisi olan yapay sinir aęları teknięi ve lojistik regresyon analizleridir. LRA da, baęımsız deęişkenlerin aldığı deęerlerle sınıflayıcı ya da sıralayıcı ölçek yapısında olan birimlerin, baęımlı deęişkene göre sınıflandırılması yapılabilmektedir. Ayrıca, LRA yardımıyla, baęımsız deęişkenlerle, sonuç (baęımlı) deęişkeni arasındaki ilişki risk yönünden incelenebilmektedir (16). Genel olarak YSA da ise, beynin bir işlevini yerine getirme yöntemini modellemek için tasarlanan matematiksel bir sistem yardımıyla oluşturulan model üzerinden sınıflama işlemi yapılmaktadır (21).

Sınıflama amacıyla LRA ve YSA kullanarak, belirlenen doęru sınıflandırma oranlarının karşılaştırılması amaçlanmıştır. Bunun içinde, Ercan ve arkadaşları (1) tarafından yapılan çalışmanın veri setinden çalışmamıza uygun 140 deneęe ait veriler alınmıştır. Çalışmamızda LRA analizi için baęımlı deęişken hasta memnuniyeti olarak alınmış, baęımsız deęişkenler ise gelir düzeyi, eğitim durumu, yaş, hekime gitme sayısı, odada birlikte kalınan hasta sayısı, hastanın en son yatışında hastanede kaldığı gün sayısı, en son yatarak tedavi olunan zaman süreci olarak belirlenmiştir. LRA için belirlenen baęımsız deęişkenler YSA için girdi olarak kabul edilmiş ve çok katmanlı yapay sinir aęı modeli buna göre oluşturulmuş ve sınıflama yapmak amacıyla oluşturulmuş model bu deęişkenlere göre belirlenmiştir.

140 kişiden toplanan verilere LRA analizi uygulandığında üç deęişkenin sonuca anlamlı etkisi olduğu belirlenmiştir. Bunlar eğitim durumu, odada beraber kalınan hasta sayısı ve hastanın en son yatışında hastanede kaldığı gün sayısı deęişkenleridir.

Eğitim durumu için ilköğretim mezunu olmak referans olarak alındığında dięer grupların riskleri sırasıyla 3.25 ve 1.43 tür.

Odada beraber kalınan hasta sayısı için 2 ve daha az kişi sayısı referans alındığında dięer grup için risk 2.8 olarak bulunmuştur.

Hastanın son yatışında hastanede kaldığı gün sayısı için 3 ve daha az gün yatmış olmak referans olarak alındığında dięer grup için risk 2.73 olarak bulunmuştur.

YSA için ise LRA için kullanılan baęımsız deęişkenler girdi olarak kullanılmış, uygulamada kullanılan veri setinin sınıflandırılması için girdi katmanında 7 birim, çıktı katmanında 1 birim olan ve iki gizli katmana sahip, ilk gizli katmanında 11 ikinci gizli katmanında 10 birim olan üç katmanlı bir yapay sinir aęı modeli kurulmuştur. Veri seti,

%30'u eğitim, %30'u geçerlilik ve %40'ı test seti olmak üzere üçe ayrılmıştır. En yüksek sınıflandırma oranını elde edebilmek için gizli katmanlar ve sahip oldukları birim sayıları simülasyon yapılarak 1000 tane yapay sinir ağı modeli denenmiştir. Gizli katman ve çıktı katmanında aktivasyon fonksiyonu olarak sigmoid lojistik fonksiyonu kullanılmıştır.

Chaseevuk ve arkadaşları (22) danışmanlı öğrenme paradigmalarından en iyi bilinen üç tanesinin (geriye yayımlı, radyal tabanlı ve öğrenme vektörü) sınıflama özelliklerini simülasyon çalışması yaparak LRA ile karşılaştırmışlar sonuçta: geriye yayılım algoritması kullanan YSA modeli için sınıflandırma oranını %56.7, radyal tabanlı fonksiyon kullanan YSA modeli için doğru sınıflandırma oranını %23.3, öğrenme vektörü algoritması kullanan YSA modeli için doğru sınıflandırma oranını %36.7 ve lojistik regresyon için doğru sınıflandırma oranını % 53.3 olarak bulmuşlardır. Bartfay ve arkadaşları (23) aynı veri seti üzerinde YSA ve LRA kullanarak yaptıkları çalışmada doğru sınıflandırma oranlarını karşılaştırmayı amaçlamışlar beş farklı lojistik regresyon ve üç farklı YSA modeli arasından doğru sınıflandırma oranları en iyi olan LRA ve YSA modellerini almışlar ve LRA için doğru sınıflandırma oranını %65, YSA için doğru sınıflandırma oranını ise %67 olarak hesaplamışlardır.

Çalışmamızda, 140 hastadan elde edilen verilere LRA uygulandığında verilerin doğru sınıflandırma oranı 7 değişken için % 81.4 olarak bulunmuştur. YSA uygulandığında ise doğru sınıflandırma oranı % 85 olarak bulunmuştur.

LRA, bağımlı değişken üzerinde etkisi olan değişkenleri modele alarak etkili olmayan değişkenleri eleyebilme yeteneğine sahiptir. YSA ise etkisi olmayan değişkenleri modelden çıkaramamaktadır. LRA sadece lojistik fonksiyonu kullanan istatistiksel modelleme yöntemidir.

YSA' da ise araştırmacı, amacına göre farklı YSA oluşturma, kullanacağı yapay sinir ağını kendisi tasarlayabilme ve farklı tasarımları deneyerek veriye en uygun ağı oluşturabilme özgürlüğüne sahiptir.

LRA' da veri setini ikiye ayırma zorunluluğu yoktur, YSA' da ise aşırı uyumdan kaçınmak için durdurma kriteri olarak verilerin en az % 10'unun test setine ayrılması gerekir. Bu durumda veri setinin hacminin azalmasına neden olur.

LRA ile uygun modelin bulunması bilgisayarla çok kısa zamanda gerçekleşebilir. YSA' da veriye en uygun modele karar verilmesi süreci çok fazla bilgisayar zamanı gerektirmektedir. LRA ile oluşturulan modelin yapısı yapay sinir ağına göre daha basittir. LRA' da, modelin parametre tahminleri odds oranları hakkında detaylı bilgi verir.

YSA, modelin gizli tabaka ve gizli nöron sayısı fazla olduğunda parametre tahminlerinin yorumlanması zorlaşır ve aynı zamanda odds oranları hakkında detaylı bilgi vermez.

Bu sonuçlara göre, YSA modelleri ile sınıflandırmanın LRA kullanılarak yapılan sınıflandırmadan daha iyi sonuçlar verme eğiliminde olduğu ayrıca yine aşırı eğitime, mimarinin hatalı oluşturulması vb. problemleri olmayan YSA modellerinin daha iyi öngörü performansı sağlayabildiği görülmektedir (24). YSA, farklı istatistiksel işlemlerin yerine getirilmesinde önemli bir rol oynamaktadır. Doğrusal olan ve doğrusal olmayan modellerin seçim imkânını tanıması YSA' nın LRA' ya göre bir üstünlüğüdür (25). Buda istatistiksel modellerde, tahmin analizlerinde, sınıflandırma problemlerinde YSA' nın önemini yükseltmektedir.

KAYNAKLAR

1. ERCAN İ, EDİZ B, HACIMUSTAFAOĞLU M, KAN İ, BOSTAN Ö. Kümeleme Çözümlemesinin Yeni Doğan Sarılıklı Olgulara Uygulanması, Uludağ Üni. Tıp Fak. Derg. Vol 24 (1-2-3) : 17-22, 1997.
2. LAWANGA SK, TYE CY. Sağlık İstatistiği Öğretimi: Yirmi Model Ders ve Seminer, çev. : Özgür S. Törel S. Cumhuriyet Üni. Yayınları No: 31, Ankara, 1990, 249, 260.
3. SÜMBÜLOĞLU K, SÜMBÜLOĞLU V. Biyoistatistik, Hatipoğlu Yayınevi, Ankara, 1990.
4. KAN İ. Biyoistatistik, II Basım, Uludağ Üni. Güçlendirme Vakfı Yayını No: 93, Bursa, 1994.
5. ÖZDAMAR K. Hastalık Olgularının İncelenmesinde Kümeleme Çözümlemesinin Kullanılması, Anadolu Üni. Tıp Fak. Yayınları No 25, Eskişehir 1998, 1-39.
6. TATLIDİL H. Uygulamalı Çok Değişkenli İstatistiksel Analiz, Engin Yayınları, Ankara, 1992, 11, 122, 252-260.
7. HOSMER D W, LEMESHOW S. Applied Logistic Regression. John Wiley & Sons, 1998.
8. ROSNER B. Multivariate Methods in Ophthalmology with Application to Other Paired Data Situations. Biometrics, 40: 1025-1035, 1984.
9. KAHN H A, SEMPAS C T. Statistical Methods in Epidemiology.
10. ELMAS Ç. Yapay Sinir Ağları, Seçkin Yayıncılık, Ankara, 2003.
11. MEMMEDOV M, ERYILMAZ H. Yapay Sinir Ağları İle Bazı İstatistiksel Modeller Arasındaki İlişki, International XII. Turkish Symposium on Artificial Intelligence and Neural Network, 2003.
12. FAUSETT L V. Fundamentals Of Neural Networks, Printice-Hall, Inc., New Jersey, (1994).

13. AKPINAR H. Yapay Sinir Ağları Ve Kredi Taleplerinin Değerlendirilmesinde Bir Uygulama Önerisi, İstanbul Üniversitesi İşletme Fakültesi Sayısal Yöntemler Anabilim Dalı, İstanbul, 1993.
14. BISHOP C.M. Neural Networks For Pattern Recognition, Clarendon Press, Oxford, 1995.
15. WANG, S. *An Adaptive Approach To Market Development Forecasting*, Neural Computing & Applications 8, 3-8, 1999.
16. EDİZ B. Lojistik Regresyon-Ayırma Analizi, Ayrımsama Sorunu ve Kalp Hastalarında Lojistik Model Yardımıyla Risk Ölçütlerinin Belirlenmesi, Doktora Tezi, Uludağ Üniversitesi Sağlık Bilimleri Enstitüsü, Bursa, 1997.
17. YARANDI H Y. Coddling Dummy Variables and Calculating the Relative Risk in a Logistic Regression , Nurs. Res., 42(5):312-314, 1993.
18. CHRISTIENSEN R. Loglinear Models, Springer-Verlag, 1990
19. GÜNERİ N. Öğrenci Başarısızlıklarının Analizinde Sinir Ağları Yaklaşımının Lojistik Regresyon Analizi ile Karşılaştırılması, Yüksek Lisans Tezi, Ankara Üni. Fen Bilimleri Enstitüsü, Ankara, 2001.
20. ADIGÜZEL F. Neural Networks As A Statistical Tool, Master Thesis, Middle East Technical University, Ankara, 2001.
21. <http://www.backpropagation.netfirms.com>
22. CHAVEESUK et al. Alternative Neural Network Approaches to Corporate Bond Rating, Engineering Valuation and Computational Intelligence of Journal of Engineering Valuation and Cost Analysis, vol. 2, no. 2, 1999, 117-131.
23. BARTFAY et al. Comparing the predictive value of neural network models to logistic regression models on the risk of death for small-cell lung cancer patients, (2006) European Journal of Cancer Care 15 , 115–124.
24. YURTOĞLU H. Yapay Sinir Ağları Metodolojisi İle Öngörü Modellemesi: Bazı Makroekonomik Değişkenler İçin Türkiye Örneği, Uzmanlık Tezi, DPT – Uzmanlık Tezleri, Yayın No: DPT:2683, 2005

25. JAÏMES et al. Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room. *Critical Care* 2005, 9:R150-R156.

TEŞEKKÜR

Yüksek lisans tezimi gerçekleştirmemde bana yardımcı olan değerli hocam Prof. Dr. İsmet KAN 'a teşekkürlerimi bir borç bilirim.

- Yüksek lisans tezimi gerçekleştirmemdeki desteklerinden dolayı değerli hocalarım Yard. Doç. Dr. Bülent EDİZ ve Dr. İlker ERCAN'a ,
- Değerli meslektaşlarım Arş. Gör. Güven ÖZKAYA'ya, Arş. Gör. Şenay YOLAÇAN'a ve Arş. Gör. Deniz SİĞİRLİ'ya ,
- Tezimle ilgili resmi işlemlerde bana yardımcı olan tüm Sağlık Bilimleri Enstitüsü personeline,
- Sağlık Bilimleri Enstitüsü emekli personeli Yusuf AKGÜN'e,
- Manevi desteklerinden dolayı babam Tefvik OCAKOĞLU' na, annem Nesrin OCAKOĞLU' na teşekkür ederim.

ÖZGEÇMİŞ

05 Mart 1980 tarihinde İzmir’de doğdum. İlkokulu İzmir Mehmet Akif Ersoy ilkokulunda bitirdim. 1992 yılında Güzelyalı Ortaokuluna kayıt yaptırđım ve orta öğrenimimi yine bu okulda tamamladım. 1995 yılında İzmir Atatürk Süper Lisesini kazandım ve lise öğrenimimi yine bu okulda tamamladım. 1999 yılında Kırıkkale Üniversitesi Fen-Edebiyat Fakültesi İstatistik bölümünü kazandım. 2000 yılında Eskişehir Anadolu Üniversitesi Fen-Edebiyat Fakültesi İstatistik bölümüne yatay geçiş yaptım ve 2003 yılında yine aynı üniversiteden mezun oldum.