

ÇOK KATMANLI PERSEPTRON SİNİR AĞLARI İLE DİYABET HASTALIĞININ TEŞHİSİ

İnan GÜLER ve Elif Derya ÜBEYLİ*

Elektronik-Bilgisayar Bölümü, Teknik Eğitim Fakültesi, Gazi Üniversitesi, Teknikokullar, 06500 Ankara,
iguler@gazi.edu.tr

* Elektrik-Elektronik Mühendisliği, Mühendislik Fakültesi, TOBB Ekonomi ve Teknoloji Üniversitesi,
Söğütözü, 06530 Ankara, edubevli@etu.edu.tr

(Geliş/Received: 22.02.2005; Kabul/Accepted: 04.07.2005)

ÖZET

Yapay sinir ağları (YSA), farklı disiplinlerdeki karmaşık problemlerin çözülmesinde kabul gören ve uygulamalarda sıklıkla yer alan modelleme araçları haline gelmiştir. Farklı YSA yapıları, tıp alanında karar destek sistemlerinin gelişmesinde kullanılmakta olan önemli modellerdendir. Bu çalışmada, dört farklı algoritma ile eğitilen çok katmanlı perseptron sinir ağları diyabet hastalığının teşhisinde kullanılmış ve en başarılı algoritma belirlenmiştir. Geri yayılım, delta-bar-delta, genişletilmiş delta-bar-delta ve hızlı yayılım, çalışılmış olan dört algoritmadır. Çok katmanlı perseptron sinir ağlarının eğitimi, geçerliliği ve testi veri tabanında yer alan farklı kişilere ait kayıtlar ile yapılmıştır. Performans belirleyiciler ve istatistiksel ölçümler ile çok katmanlı perseptron sinir ağları değerlendirilmiş ve sonuçlar diyabet hastalığının teşhisinde hızlı yayılım algoritmasının en başarılı çok katmanlı perseptron eğitim algoritması olduğunu göstermiştir.

Anahtar Kelimeler: Çok katmanlı perseptron sinir ağı, eğitim algoritmaları, diyabet hastalığının teşhisi.

DIABETES DIAGNOSIS BY MULTILAYER PERCEPTRON NEURAL NETWORKS

ABSTRACT

Artificial neural networks (ANNs) have become modeling tools that have found extensive acceptance and they have frequently used in applications in many disciplines for solving complex problems. Different ANN structures are valuable models, which are used in the medical field for the development of decision support systems. In this study, four multilayer perceptron neural networks (MLPNNs) trained with different algorithms were used for diabetes prediction and the most efficient training algorithm was determined. Backpropagation, delta-bar-delta, extended delta-bar-delta and quick propagation were the studied four training algorithms. The MLPNNs were trained, cross validated and tested with subject records from the database. Performance indicators and statistical measures were used for evaluating the MLPNNs and the results demonstrated that the quick propagation algorithm was the most efficient multilayer perceptron training algorithm for diabetes prediction.

Keywords: Multilayer perceptron neural network, training algorithms, diabetes diagnosis.

1. GİRİŞ

Tıbbi karar verme çok yönlü bir işlemdir ve amaç doğru teşhis yapılmasıdır. Bu amaca ulaşmak için uygun verinin bulunması, verinin özelliklerinin çıkarılması ve yeni verinin analizinin yapılması gerekmektedir. Hekimler, karar verme işleminde çeşitli istatistiksel teknikler ile veriyi işlemektedirler. Verinin karmaşıklığı ve boyutunun artması durumunda veri analizi için bilgisayarların kullanımı

gereklidir. İstatistiksel analizlerin bilgisayar ile yapılmasının yanı sıra bilgisayar destekli veri sınıflama, yapay sinir ağları (YSA) uygulamaları arasında yer almaktadır. Son zamanlardaki gelişmeler incelendiğinde, tıpta bilgisayar destekli karar verme uygulamalarına yönelik çalışmalarda artış olduğu görülmektedir. Hastalık teşhisleri, veri sınıflama işlemi olarak incelenebilmektedir. Girişlerin belirsiz olması ve değişkenlik göstermesi durumunda YSA veri sınıflamada başarılı olmaktadır [1-4].

YSA girişleri (sistemin bağımsız değişkenleri) çıkışlar (bağımlı kestirilen değişkenler) ile ilişkilendirerek karmaşık, doğrusal olmayan modeller oluşturur. YSA'nın modelleme ve karar verme işlemlerinde kullanılan en önemli araçlardan olmasının nedeni, sınırlı ve tamamlanmamış olan veri tabanlarından en iyi sonucu çıkarabilmesi ve farklı eğitim algoritmaları ile eğitilmeleri durumunda başarının artırılabilmesidir. Bu özelliklerinin yanı sıra, YSA'nın klinik protokollerden çıkarılan veriler, ölçümlerden elde edilen laboratuvar verileri, işaretler veya görüntülerden çıkarılan öznitelikler gibi bir sistemin farklı yapıdaki verilerini birleştirerek tümleşik teşhis sistemi oluşturma özelliği vardır [2-4].

Farklı disiplinlerdeki problemlerin çözülmesinde kullanılabilen YSA için farklı ağ mimarileri ve farklı eğitim algoritmaları geliştirilmiştir. YSA'nın uygulamalarında önemli bir yeri olan çok katmanlı perseptron sinir ağları, birçok sezim ve kestirim işlemlerini yürütmek için kullanılan parametrik olmayan bir tekniktir. Çok katmanlı perseptron sinir ağlarının eğitimi için farklı eğitim algoritmaları kullanılabilir. Bu algoritmalar en çok kullanılmakta olan geri yayılım algoritmasıdır ve bu algoritmanın bazı problemleri vardır [5-7]. Bu problemleri en aza indirebilmek için delta-bar-delta, genişletilmiş delta-bar-delta ve hızlı yayılım algoritmaları geliştirilmiştir [8-10]. Literatürde çok katmanlı perseptron sinir ağlarının eğitiminde kullanılan algoritmaların performansını karşılaştıran çalışmalar yer almaktadır [11-13]. Bu çalışmaların sonuçları, algoritmaların performanslarının üzerinde çalışılan probleme bağlı olarak değiştiğini ortaya çıkarmıştır. Bu durumda, algoritmaların performansları ile ilgili genel bir sonuç belirtilememektedir.

Bu çalışmada, diyabet hastalığının teşhisi için çok katmanlı perseptron sinir ağları dört farklı algoritma ile eğitilmiştir. Geri yayılım ve geri yayılım algoritması ile ilişkisi tanımlı olan delta-bar-delta, genişletilmiş delta-bar-delta ve hızlı yayılım algoritmaları çalışılmış olan dört algoritmadır. Çok katmanlı perseptron sinir ağlarının geliştirilmesinde MATLAB program paketi kullanılmıştır. Performans belirleyiciler ve istatistiksel ölçümler ile geliştirilmiş olan çok katmanlı perseptron sinir ağlarının değerlendirilmesi yapılmış ve diyabet hastalığının teşhisi için en yüksek performansı gösteren çok katmanlı perseptron eğitim algoritması belirlenmiştir.

2. DİYABET HASTALIĞININ VERİ TABANI VE ÇOK KATMANLI PERSEPTRON SİNİR AĞLARININ GİRİŞ PARAMETRELERİ

Diyabet hastalığının veri tabanında 760 kişiye ait kayıtlar yer almaktadır [14]. Grup, yaşları 10 ile 75 arasında değişen 412 kadın ve 348 erkekten oluşmaktadır. Muayene sonuçları, 760 kişiden 268'inin diyabet hastası geriye kalanların ise sağlıklı olduğunu göstermiştir. Diyabet hastası olanların 151'i kadın 117'si erkek (yaş aralığı 10-75), sağlıklı olanların 261'i kadın 231'i erkektir (yaş aralığı 10-71).

Diyabet, pankreas tarafından salgılanan insulinin eksikliği veya yokluğu durumunda gözlenen bir metabolizma hastalığıdır. Diyabet iki ana formda oluşur: Tip I, veya insulin bağımlı diyabet ve Tip II veya insulin bağımsız diyabet. Bu çalışmada kullanılan veri tabanında özellikle kişilerin yaşlarının 10 ile 75 arasında olmasından dolayı hem Tip I hem Tip II incelenmektedir. Diyabet hastalığının önemli risk faktörlerinden oluşan ondört bağımsız parametre çok katmanlı perseptron sinir ağlarının girişi olarak kullanılmıştır. Ağ girişi olarak kullanılan bu ondört parametre Tablo 1'de verilmektedir. Tip I diyabeti çoğunlukla çocuklarda ve otuz yaş altı olan yetişkinlerde görülmektedir. Çocuk ve genç yetişkinler arasında kadın ve erkeklerdeki yaş dağılımı çok küçük farklılıklar göstermektedir. Hem kadınlar hem erkekler hastalıktan etkilenmektedir fakat birçok toplulukta Tip II diyabeti çoğunlukla kadınlarda gözlenmektedir. Tip II diyabetinin oluşum riski yaşla birlikte artmaktadır ve birçok toplulukta bu tip diyabet orta yaş ve orta yaş üstü olan kişilerde görülmektedir. Plazma glikoz konsantrasyonunun normal düzeyi 80-110mg/dl'dir. Diyabet hastalarında açlık plazma glikoz konsantrasyonu 126mg/dl'nin üzerinde iken yemekten 2 saat sonraki plazma glikoz konsantrasyonu 200mg/dl'nin üzerindedir. Tip I diyabetinde beta hücre hasarına bağlı mutlak insulin eksikliği görülmektedir ve pankreasta insulin içeriği 0µU/ml değerine düşer. Tip II diyabetinde ise insulin direnci bulunabilir. Bu tip diyabette insulin sensitivitesi azalmıştır ve pankreatik insulin içeriği normal değerin %50'sinden fazladır. Diyastolik kan basıncının 90mmHg değerinden yüksek olması diyabet hastalığı için bir risk faktörüdür. Diyabet hastalarının vücutlarındaki bütün damarlar anormalleşir. Diyabet hastalarında görülen deri hastalıklarının nedeni damar anormallikleridir. İnsulin tedavisi refraktif değişimlere neden olduğu için diyabet hastalarında görme bozuklukları oluşabilir. Vücut kütle oranı diyabet gibi birçok hastalığın risk faktörü olarak incelenir. Vücut kütle oranı, 18.5'ten küçük zayıf, 18.5-24 aralığında normal, 25-29.9 aralığında aşırı kilolu, 30'dan yukarısı obez olmak üzere dört durumda ele alınır. Vücut kütle oranı diyet ile kontrol edilebilir. Diyabet düzensiz beslenme ile ilişkili olarak ortaya çıkabilir. Beslenme ve diyabet arasındaki ilişki iki ayrı yönden incelenmelidir: ilk olarak beslenmenin diyabeti oluşturmaya ve ikinci olarak insulin bağımlı veya insulin bağımsız diyabet hastalarına uygun olan besin çeşitleri. Kolesterolün 239mg/dl değerinden fazla olması diyabet risk faktörlerinden biridir. Fiziksel aktiviteler diyabet tedavisinde ve vücut kütle oranlarının ayarlanmasında önemlidir. Araştırma sonuçları, fiziksel aktivitenin diyabet hastalığının oluşmasını engelleme olasılığı olduğunu göstermektedir. Uyku olma diyabet belirtilerinden biridir ve diyabet hastaları için günlük uyku saati 9 saatten daha fazladır. Kalıtsal olmasından veya aynı aileden olan kişilerin ortak ortamı paylaşımlarından dolayı diyabet aile içerisindeki diğer kişilerde görülebilmektedir. Bundan dolayı, ailedeki kişilerde diyabet

Tablo 1. Çok katmanlı perseptron sinir ağlarının giriş parametreleri

İsim	Tanım	Tip	Değerler	
			Ortalama	Standard sapma
Cinsiyet	Kadın/Erkek	Sınıflandırılabilir 0/1	—	—
Yaş	Kişilerin yaşı	Sayısal	33.2	11.8
Plazma glikoz	Plazma glikoz konsantrasyonu (mg/dl)	Sayısal	120.9	32.0
Plazma insulin	Plazma insulin konsantrasyonu (μ U/ml)	Sayısal	79.8	115.2
Diyastolik kan basıncı	Diyastolik kan basıncı (mmHg)	Sayısal	69.1	19.4
Deri hastalığı	Deri hastalığının görülmesi (Hayır/Evet)	Sınıflandırılabilir 0/1	—	—
Bulanık görme	Bulanık görme (Hayır/Evet)	Sınıflandırılabilir 0/1	—	—
Vücut kütle oranı	Ağırlık (kg) / (Boy (m)) ²	Sayısal	28.5	8.2
Diyet	Diyet (Hayır/Evet)	Sınıflandırılabilir 0/1	—	—
Kolesterol	Kolesterol (mg/dl)	Sayısal	224.6	20.1
Fiziksel aktivite	Fiziksel aktivite (Hayır/Evet)	Sınıflandırılabilir 0/1	—	—
Uyku saatleri	Bir gündeki uyku saati	Sayısal	7.2	5.6
Kahtım	Ailede diyabet hastalığının görülmesi (Hayır/Evet)	Sınıflandırılabilir 0/1	—	—
Gebelik	Gebelik sayısı	Sayısal	3.8	3.4

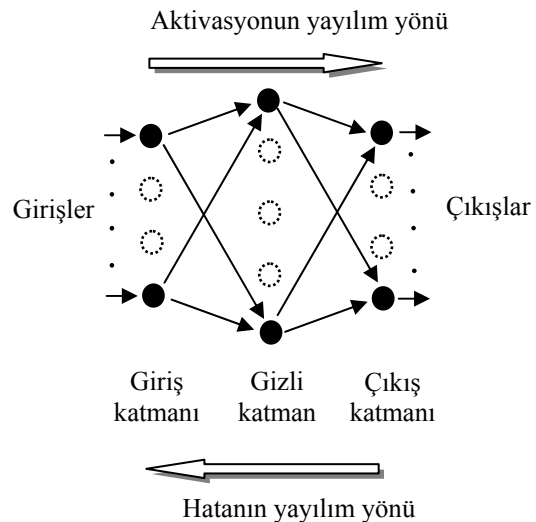
hastalığının olup olmadığı belirlenmelidir. Gebelik sırasında oluşan veya ilk olarak gebelik sürecinde belirlenen gebelikteki diyabet, daha sonraki yıllarda diyabet hastalığının oluşma riskini artırmaktadır. Bundan dolayı, gebelikte diyabet hastası olan kadınlarda Tip II diyabeti görülme olasılığı oldukça yüksektir [15-20].

3. ÇOK KATMANLI PERSEPTRON EĞİTİM ALGORİTMALARI

Çok katmanlı perseptron sinir ağları, giriş nöronlarının bulunduğu giriş katmanı, çıkış nöronlarının bulunduğu çıkış katmanı ve bir veya daha fazla gizli katmandan oluşur. Doğrusal olmayan problemlerin çözümlenmesinde giriş katmanı ile çıkış katmanı arasında ilave katman(lar) yer aldığı için ağ mimarisi Şekil 1’de görüldüğü gibi çok katmanlı olmaktadır. Sinir ağlarının mimarisinde gizli katman sayısının belirlenmesi oldukça önemlidir. Giriş ve çıkış katmanlarının oluşturulmasından farklı olarak ağ mimarisi gerçekleştirilirken gizli katman sayısı hakkında ön bilgi olmadan işleme başlanır. Birkaç gizli nöronu olan ağ karmaşık örüntüleri ayırt edememektedir çünkü sadece doğrusal kestirim yapabilmektedir. Bunun yanı sıra, gizli nöron sayısının fazla olması ağın genelleştirme yapmasını engellemektedir. Gizli katman sayısının artırılması durumunda ağın eğitimi için geçen süre artmaktadır. Genellikle ağ mimarilerinin oluşturulmasında, deneme yanılma yöntemi ile uygun olan gizli katman sayısı tespit edilir [5,6]. Bu çalışmada yer alan çok katmanlı perseptron ağ mimarisinde gizli katman sayısı ikidir ve deneme yanılma yöntemi ile belirlenmiştir. Ağ mimarisinin oluşturulmasında katmanlardaki nöron sayıları da

deneme yanılma yöntemine göre tespit edilmiştir. Yapılan testler sonucu en yüksek performansı gösteren ağ mimarisi $14 \times 10 \times 15 \times 2$ (giriş nöronları $\times$ gizli katman 1’in nöronları $\times$ gizli katman 2’nin nöronları $\times$ çıkış nöronları) olarak bulunmuştur.

YSA’nda bilgi nöronlar arasındaki bağlantıların ağırlıklarında yer almaktadır. Bundan dolayı, bilgiye ait olan temsilci örneklerin ardışık olarak ağı verildiği ve ağın bu bilgiyi kendi yapısı içerisinde bütünleştirmesini sağlayan eğitim işlemi YSA uygulamalarında oldukça önemlidir. Bu çalışmada, çok katmanlı perseptron sinir ağlarının eğitiminde kullanılan dört algoritma burada kısaca verilmektedir [5,6].

**Şekil 1.** Çok katmanlı perseptron sinir ağı

Çok katmanlı perseptronda gizli katmandaki j indisli her bir nöron, x_i giriş işaretini bağlantının ağırlığı w_{ji} ile çarptıktan sonra toplar ve y_j çıkışını toplamın fonksiyonu olarak şu şekilde hesaplar:

$$y_j = f\left(\sum w_{ji}x_i\right) \quad (1)$$

Burada, f aktivasyon fonksiyonudur ve nöronlardaki işaretlerin ağırlıklı toplamalarını dönüştürmek için gerekmektedir. Bu çalışmada, giriş, gizli ve çıkış katmanlarında aktivasyon fonksiyonu olarak sıfır ile bir aralığında değişen ve iki önemli özelliği olan sigmoid fonksiyonu kullanılmıştır. Sigmoid fonksiyonu doğrusal değildir ve ağ girişlerinin çıkış vektör uzaylarına karmaşık dönüşümünü yapabilmektedir. Ayrıca, sürekli bir fonksiyondur ve türevi alınabilmektedir. Böylece, ağırlıkların ayarlanmasında kullanılan hatanın değişim ölçüsü elde edilebilmektedir. Çıkış nöronunun istenilen değeri ve gerçek değeri arasındaki farkın toplamının karesi, E , şu şekilde ifade edilir:

$$E = \frac{1}{2} \sum_j (y_{dj} - y_j)^2 \quad (2)$$

Burada, y_{dj} j indisli çıkış nöronunun istenilen değeri ve y_j bu çıkış nöronunun gerçek değeridir. Çok katmanlı perseptron sinir ağı, ağınc çıkışındaki hatayı hesaplar ve hatayı azaltmak için nöronların ağırlıklarını yeniden düzenler. Her bir w_{ji} ağırlığının yeniden düzenlenmesinde ağırlıklara Δw_{ji} eklenir. Δw_{ji} 'nin hesaplanması eğitim algoritmasına bağlıdır [5,6].

3.1. Geri Yayılım Algoritması

Rumelhart ve arkadaşları [7] tarafından sunulan geri yayılım algoritmasına göre k 'nci iterasyondaki i ve j indisli nöronlar arasında mevcut olan bağlantıdaki değişim $\Delta w_{ji}(k)$ şu şekilde ifade edilir:

$$\Delta w_{ji}(k) = -\alpha \frac{\partial E}{\partial w_{ji}(k)} + \mu \Delta w_{ji}(k-1) \quad (3)$$

Burada, α öğrenme katsayısı, μ momentum katsayısı ve $\Delta w_{ji}(k-1)$ bir önceki iterasyondaki ağırlık değişimidir. Öğrenme katsayısı, eğitim sırasında her bir ağırlığa yapılan ayarlama ile orantılıdır ve ağırlıkların çözüme yaklaşma hızını etkiler. Momentum katsayısı, her iterasyonda hesaplanan ağırlık hatasının türevi ile orantılıdır ve uygun değerleri geri yayılım algoritmasının doğru olan minimuma yaklaşmasını sağlar. Öğrenme ve momentum katsayılarının uygun değerleri ile çözüm için hızlı ve kararlı yaklaşımların elde edilmesi mümkün olmaktadır. Bununla birlikte, geri yayılım algoritmasının birçok uygulamada bazı problemleri vardır. Geri yayılım algoritması ile

eğitilen çok katmanlı perseptron sinir ağlarının en önemli problemlerinden biri, kabul edilebilir ağırlık çözümüne ulaşılabilmesi için uzun eğitim süresi geçmektedir. Delta-bar-delta, genişletilmiş delta-bar-delta ve hızlı yayılım eğitim algoritmaları, çok katmanlı perseptron sinir ağlarının eğitimini hızlandırmak için geliştirilmiş olan eğitim algoritmalarıdır.

3.2. Delta-Bar-Delta Algoritması

Delta-bar-delta eğitim algoritması ağıdaki bütün ağırlıkları ayarlamak için kullanılır ve k 'nci iterasyondaki i ve j indisli nöronlar arasında mevcut olan bağlantıdaki değişim $\Delta w_{ji}(k)$ şu şekilde ifade edilir [8]:

$$\Delta w_{ji}(k) = -\alpha_{ji}(k) \frac{\partial E}{\partial w_{ji}(k)} \quad (4)$$

Burada $\alpha_{ji}(k)$ i indisli nöronun j indisli nörona olan bağlantının öğrenme katsayısıdır. Öğrenme katsayısındaki değişim şu şekildedir:

$$\Delta \alpha_{ji}(k) = \begin{cases} A, & D_{ji}(k-1) \frac{\partial E}{\partial w_{ji}(k)} > 0 \\ -\varphi \alpha_{ji}(k), & D_{ji}(k-1) \frac{\partial E}{\partial w_{ji}(k)} < 0 \\ 0, & \text{diğer durumlar} \end{cases} \quad (5)$$

Burada $D_{ji}(k-1)$, $\partial E / (\partial w_{ji}(k-1))$ ve $\partial E / (\partial w_{ji}(k-2))$ ifadelerinin ağırlıklı ortalamasını göstermektedir ve şu şekilde belirtilir:

$$D_{ji}(k-1) = (1-\theta) \frac{\partial E}{\partial w_{ji}(k-1)} + \theta \frac{\partial E}{\partial w_{ji}(k-2)} \quad (6)$$

ve φ , θ , A pozitif sabitlerdir.

Eşitlik 5'ten görüldüğü gibi, algoritma öğrenme katsayılarını doğrusal olarak artırmakta ve geometrik olarak azaltmaktadır.

3.3. Genişletilmiş Delta-Bar-Delta Algoritması

Algoritmanın isminden de anlaşıldığı üzere bu algoritma delta-bar-delta algoritmasının genişletilmiş şeklidir [9]. Ağırlıklardaki değişim şu şekilde hesaplanır:

$$\Delta w_{jk}(k) = -\alpha_{ji}(k) \frac{\partial E(k)}{\partial w_{ji}(k)} + \mu_{ji}(k) \Delta w_{ji}(k-1) \quad (7)$$

Burada $\alpha_{ji}(k)$ öğrenme katsayısı, $\mu_{ji}(k)$ momentum katsayısıdır.

Genişletilmiş delta-bar-delta algoritması ile delta-bar-delta algoritması arasındaki farklardan biri, genişletil-

miş delta-bar-delta algoritmasında momentum katsayısının kullanılıyor olmasıdır. Öğrenme katsayısındaki değişim şu şekildedir:

$$\Delta\alpha_{ji}(k) = \begin{cases} A_\alpha \exp(-\gamma_\alpha |D_{ji}(k)|), & D_{ji}(k-1) \frac{\partial E}{\partial w_{ji}(k)} > 0 \\ -\varphi_\alpha \alpha_{ji}(k), & D_{ji}(k-1) \frac{\partial E}{\partial w_{ji}(k)} < 0 \\ 0, & \text{diğer durumlar} \end{cases} \quad (8)$$

Burada A_α , φ_α ve γ_α pozitif sabitlerdir ve D_{ji} delta-bar-delta algoritmasındaki gibi hesaplanır.

Momentum katsayısındaki değişim şu şekildedir:

$$\Delta\mu_{ji}(k) = \begin{cases} A_\mu \exp(-\gamma_\mu |D_{ji}(k)|), & D_{ji}(k-1) \frac{\partial E}{\partial w_{ji}(k)} > 0 \\ -\varphi_\mu \mu_{ji}(k), & D_{ji}(k-1) \frac{\partial E}{\partial w_{ji}(k)} < 0 \\ 0, & \text{diğer durumlar} \end{cases} \quad (9)$$

Burada A_μ , φ_μ ve γ_μ pozitif sabitlerdir.

Eşitlik 8'den görüldüğü gibi, öğrenme katsayısındaki artış sabit değildir ve ağırlıklı ortalama $D_{ji}(k)$ ifadesinin büyüklüğünün üssel olarak azalan bir fonksiyonudur. Ağırlık değerlerindeki salınımları önlemek için $\alpha_{ji}(k)$ ve $\mu_{ji}(k)$, önceden belirlenmiş $\alpha_{\max}$ ve $\mu_{\max}$ üst sınır değerlerinden aşağıda tutulur.

3.4. Hızlı Yayılım Algoritması

Hızlı yayılım eğitim algoritması ağıdaki bütün ağırlıkları ayarlamak için kullanılır ve k 'nci iterasyondaki i ve j indisli nöronlar arasında mevcut olan bağlantıdaki değişim $\Delta w_{ji}(k)$ şu şekilde ifade edilir [10]:

$$\Delta w_{ji}(k) = -\alpha \frac{\partial E}{\partial w_{ji}(k)} + \mu \Delta w_{ji}(k-1) \quad (10)$$

Burada α öğrenme katsayısı, $\Delta w_{ji}(k-1)$ bir önceki iterasyondaki ağırlık değişimi, μ momentum katsayısıdır ve şu şekilde tanımlanır:

$$\mu = \frac{\partial E / \partial w_{ji}(k)}{(\partial E / \partial w_{ji}(k-1)) - (\partial E / \partial w_{ji}(k))} \quad (11)$$

4. UYGULAMA SONUÇLARI

Tablo 1'de verilen diyabet hastalığının risk faktörlerinden olan on dört bağımsız parametre ağ girişi olarak kullanılmıştır. Ağın çıkış vektörleri şu şekilde tanımlanmıştır:

[0 1] = sağlıklı
[1 0] = diyabet hastası

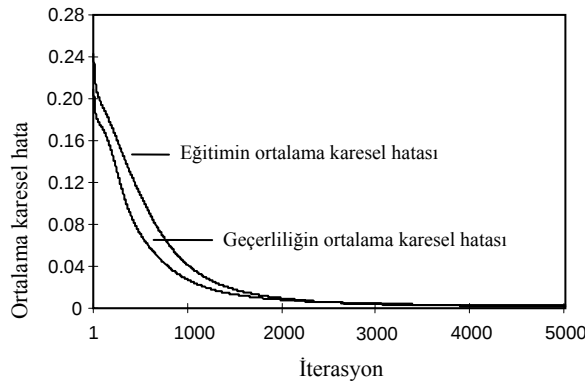
Diyabet hastalığının teşhisi için dört farklı algoritma ile eğitilen dört çok katmanlı perseptron sinir ağı geliştirilmiştir. Eğitim algoritmalarındaki parametreler deneysel olarak belirlenmiştir ve bu parametrelerin değerleri aşağıdaki gibidir:

1. Geri yayılım $\rightarrow \alpha = 0.1$, $\mu = 0.2$
2. Delta-bar-delta $\rightarrow A = 0.01$, $\varphi = 0.5$, $\theta = 0.6$
3. Genişletilmiş delta-bar-delta $\rightarrow$
 $\alpha_{\max} = \mu_{\max} = 2.0$, $A_\alpha = 0.09$, $A_\mu = 0.01$,
 $\varphi_\alpha = 0.1$, $\varphi_\mu = 0.01$, $\gamma_\alpha = 0.0$, $\gamma_\mu = 0.0$, $\theta = 0.6$
4. Hızlı yayılım $\rightarrow \alpha = 0.1$

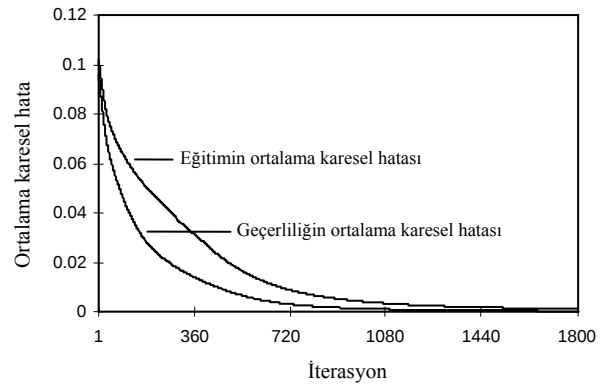
Çok katmanlı perseptron sinir ağlarının başarısı eğitim ve test için kullanılan veri boyutuna bağlı olarak değişebilmektedir. Bu çalışmada, 760 kişiden 462 kişiye ait kayıt ağın eğitimi için geriye kalanlar ise ağın testi için kullanılmıştır. Ağın daha iyi genelleştirme yapabilmesi için ağın eğitimi için kullanılan verinin küçük bir yüzdesi geçerlilik verisi olarak ele alınmıştır. Eğitim verisi olarak kullanılan 462 kişiye ait olan kayıtlardan rastgele seçilen 30 kayıt geçerlilik verisini oluşturmuştur. Eğitim verisi 295 sağlıklı, 167 diyabet hastasından; test verisi 197 sağlıklı, 101 diyabet hastasından ve geçerlilik verisi 15 sağlıklı, 15 diyabet hastasından oluşmaktadır. Çok katmanlı perseptron sinir ağlarının performans analizi, eğitim performansı ve test performansı olmak üzere iki ayrı şekilde yapılmıştır.

4.1. Çok Katmanlı Perseptron Sinir Ağlarının Eğitim Performansı

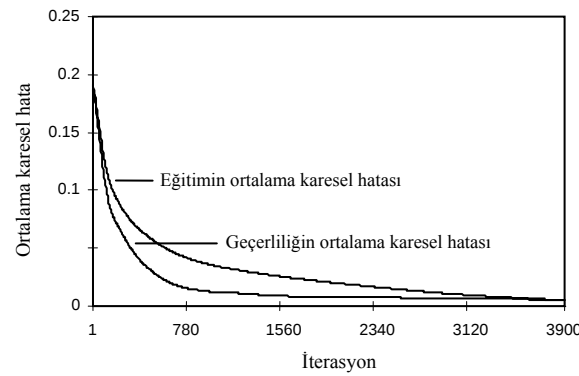
Ağın eğitimi sırasında çıkış ile istenilen sonuç arasındaki farkın azaltılması için ağırlıklar eğitim algoritmasına bağlı olarak değişir. Ortalama karesel hata eğrisi, eğitim iterasyonları boyunca ağın çıkışı ile istenilen sonuç arasındaki farkın karesini gösterir. Ağın fazla eğitilmesi durumunda ağın genelleştirme yapması zorlaşır. Ağın eğitimi durdurmak için geçerlilik kriteri kullanılır. Şekil 2-5'te dört farklı algoritma ile eğitilen ağların eğitim ve geçerlilikteki hataları aynı grafik üzerinde görülmektedir. İyi bir genelleştirme elde edilebilmesi için geçerlilikteki hatanın artması durumunda ağın eğitimi durdurulmalıdır. Algoritmaların performanslarının karşılaştırılmasında, ağların eğitimindeki merkezi işlemci zamanları incelenmiştir. Tablo 2'de her bir ağın eğitimi ve geçerliliği için minimum ortalama karesel hata, iterasyon sayısı ve ağların eğitimindeki merkezi işlemci zamanları verilmektedir. Şekil 5 ve Tablo 2'den görüldüğü gibi, hızlı yayılım algoritması ile eğitilen çok katmanlı perseptron sinir ağının ortalama karesel hatası 1800 iterasyonda sıfıra yakın bir değere ulaşmaktadır. Geri yayılım, delta-bar-delta ve genişletilmiş delta-bar-delta algoritmaları ile eğitilen çok katmanlı perseptron sinir ağların yakınsama hızları, hızlı yayılım algoritması ile eğitilen çok katmanlı perseptron sinir ağına göre oldukça düşüktür. Ağların eğitimi sırasındaki iterasyon sayıları geçerlilik kriterine göre belirlenmiştir.



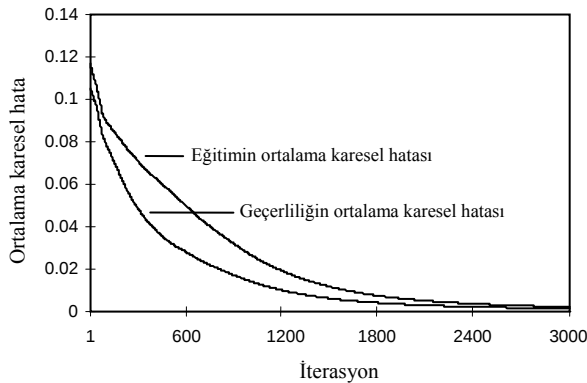
Şekil 2. Geri yayılım algoritması ile eğitilen çok katmanlı perseptron sinir ağının eğitim ve geçerlilik ortalama karesel hata eğrileri



Şekil 5. Hızlı yayılım algoritması ile eğitilen çok katmanlı perseptron sinir ağının eğitim ve geçerlilik ortalama karesel hata eğrileri



Şekil 3. Delta-bar-delta algoritması ile eğitilen çok katmanlı perseptron sinir ağının eğitim ve geçerlilik ortalama karesel hata eğrileri



Şekil 4. Genişletilmiş delta-bar-delta algoritması ile eğitilen çok katmanlı perseptron sinir ağının eğitim ve geçerlilik ortalama karesel hata eğrileri

4.2. Çok Katmanlı Perseptron Sinir Ağlarının Test Performansı

Ağın eğitiminden sonra ağın testi 298 kişiye ait olan kayıtlar ile yapılmıştır. Tablo 3'te sınıflama sonuçları verilmiştir. Tablo 3'ten görüldüğü gibi, hızlı yayılım algoritması ile eğitilen çok katmanlı perseptron sinir ağı sağlıklı kişilerden 6 kişiyi diyabet hastası olarak sınıflandırırken; 3 diyabet hastasını sağlıklı olarak sınıflandırmıştır. Tablo 3'teki sonuçlara göre hızlı yayılım algoritması ile eğitilen ağın sınıflama sonuçları diğer algoritmalar ile eğitilen ağların sınıflama sonuçlarından daha iyidir.

Çok katmanlı perseptron sinir ağlarının sınıflama performansları aşağıda verilen istatistiksel parametreler kullanılarak hesaplanmıştır:

Belirlilik: doğru sınıflandırılmış sağlıklı kişi sayısı / toplam sağlıklı kişi sayısı

Tablo 3. Sınıflama sonuçları

Farklı algoritmalar ile eğitilen ağlar	Çıkış/İstenen	Sağlıklı	Diyabet
Geri yayılım	Sağlıklı	164	16
	Diyabet	33	85
Delta-bar-delta	Sağlıklı	170	12
	Diyabet	27	89
Genişletilmiş delta-bar-delta	Sağlıklı	176	9
	Diyabet	21	92
Hızlı yayılım	Sağlıklı	191	3
	Diyabet	6	98

Tablo 2. Çok katmanlı perseptron sinir ağlarının eğitimi ve geçerliliği için minimum ortalama karesel hata, iterasyon sayısı ve ağların eğitimindeki merkezi işlemci zamanları

Farklı algoritmalar ile eğitilen ağlar	İterasyon sayısı		Minimum ortalama karesel hata		Merkezi işlemci zamanı (dak:s)
	Eğitim	Geçerlilik	Eğitim	Geçerlilik	
Geri yayılım	5000	5000	0.0065	0.0077	19:20
Delta-bar-delta	3900	3900	0.0050	0.0052	15:11
Genişletilmiş delta-bar-delta	3000	3000	0.0014	0.0022	14:03
Hızlı yayılım	1800	1800	0.0003	0.0011	8:42

Tablo 4. İstatistiksel parametreler

Farklı algoritmalar ile eğitilen ağlar	İstatistiksel parametreler		
	Belirlilik (%)	Duyarlılık (%)	Toplam doğruluk (%)
Geri yayılım	83.25	84.16	83.56
Delta-bar-delta	86.29	88.12	86.91
Genişletilmiş delta-bar-delta	89.34	91.09	89.93
Hızlı yayılım	96.95	97.03	96.98

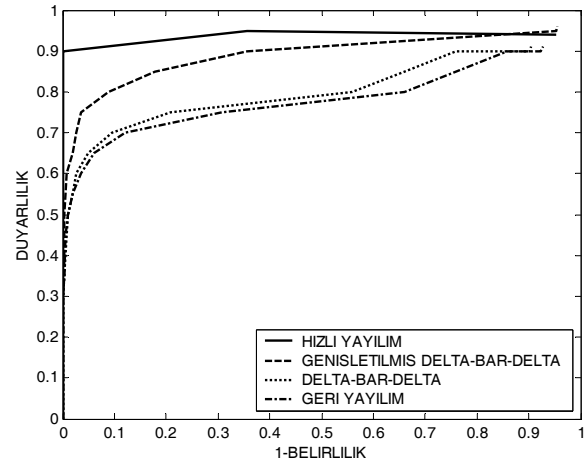
Duyarlılık: doğru sınıflandırılmış diyabet hastasının sayısı / toplam diyabet hastasının sayısı

Toplam doğruluk: doğru sınıflandırılmış kişi sayısı / toplam kişi sayısı

Her ağ için hesaplanan bu istatistiksel parametrelerin değerleri Tablo 4'te verilmiştir. Tablo 4'teki sonuçlar, hızlı yayılım algoritması ile eğitilen ağın sağlıklı kişileri %96.95 doğrulukla, diyabet hastalarını ise %97.03 doğrulukla sınıflandırdığını göstermektedir. Hızlı yayılım algoritması ile eğitilen ağın sağlıklı kişileri ve diyabet hastalarını %96.98 doğrulukla sınıflandırdığı görülmektedir. Bu sonuçlar incelendiğinde, hızlı yayılım algoritması ile eğitilen ağın sınıflama doğruluklarının diğer algoritmalar ile eğitilen ağların sınıflama doğruluklarından daha yüksek olduğu ortaya çıkmaktadır.

Çok katmanlı perseptron sinir ağlarının test performanslarının analizi için alıcı işlem karakteristik eğrileri elde edilmiştir. Alıcı işlem karakteristik eğrileri her bir ağın diyabet hastası olan kişiyi tespit etmedeki performansını göstermektedir. Sınıflama sisteminden elde edilen bir sonucun durumunu tanımlayan dört ayrı alternatif vardır: (i) doğru pozitif, (ii) yanlış pozitif, (iii) doğru negatif, (iv) yanlış negatif. Ağın pozitif olarak tespit ettiği durum hekim tarafından da pozitif olarak tespit ediliyor ise doğru pozitif kararı oluşmaktadır. Yanlış pozitif kararında, ağın pozitif olarak tespit ettiği durum hekim ile uyum sağlamamaktadır. Doğru negatif kararında, hem ağ hem hekim pozitif durum olmadığına karar vermektedir. Ağın negatif olarak tespit ettiği durum hekim ile uyum sağlamıyor ise yanlış negatif kararı oluşmaktadır. Şekil 6'da çok katmanlı perseptron sinir ağlarının test verisi üzerindeki performansını belirten alıcı işlem karakteristik eğrileri görülmektedir. Test performansı alıcı işlem karakteristik eğrilerinin şekline bağlıdır. Duyarlılığın hızla yükseldiği ve 1-belirliliğin duyarlılık yükselinceye kadar çok az artması durumunda testin başarılı olduğu söylenebilir. Şekil 6 incelendiğinde, hızlı yayılım algoritması ile eğitilen ağın test performansının diğer algoritmalar ile eğitilen ağların test performanslarından daha yüksek olduğu görülmektedir.

Ağın çıkışı ile istenilen çıkış arasındaki fark hata olarak belirtilir ve farklı yollarla ölçülebilir. Bu çalışmada, ortalama karesel hata, ortalama mutlak

**Şekil 6.** Alıcı işlem karakteristik eğrileri

hata ve korelasyon katsayısı çok katmanlı perseptron sinir ağlarının hata ölçümünde kullanılmıştır. Çok katmanlı perseptron sinir ağlarının performans değerlendirme parametreleri sağlıklı kişiler ve diyabet hastaları için Tablo 5'te verilmektedir. Tablo 5 incelendiğinde, hızlı yayılım algoritması ile eğitilen ağın ortalama karesel hata ve ortalama mutlak hatalarının diğer algoritmalar ile eğitilen ağların hatalarından daha düşük olduğu görülmektedir. Bunun yanı sıra, hızlı yayılım algoritması ile eğitilen ağın korelasyon katsayısı diğer algoritmalar ile eğitilen ağların korelasyon katsayılarına göre 1'e daha yakındır. Başarılı eğitim ve doğru ağ mimarisinden dolayı hızlı yayılım algoritması ile eğitilen ağın test performansı yüksek bulunmuştur.

Tablo 5. Performans değerlendirme parametreleri

Farklı algoritmalar ile eğitilen ağlar	Performans	Sağlıklı	Diyabet
Geri yayılım	Ortalama karesel hata	0.0098	0.0087
	Ortalama mutlak hata	0.0127	0.0149
	Korelasyon katsayısı	0.85	0.86
Delta-bar-delta	Ortalama karesel hata	0.0058	0.0063
	Ortalama mutlak hata	0.0105	0.0117
	Korelasyon katsayısı	0.87	0.86
Genişletilmiş delta-bar-delta	Ortalama karesel hata	0.0035	0.0041
	Ortalama mutlak hata	0.0102	0.0109
	Korelasyon katsayısı	0.90	0.91
Hızlı yayılım	Ortalama karesel hata	0.0026	0.0033
	Ortalama mutlak hata	0.0097	0.0101
	Korelasyon katsayısı	0.95	0.94

5. SONUÇLAR

Geri yayılım, delta-bar-delta, genişletilmiş delta-bar-delta ve hızlı yayılım algoritmaları ile eğitilen çok katmanlı perseptron sinir ağları diyabet hastalığının teşhisi için geliştirilmiştir. Hızlı yayılım algoritmasının gerçekleştirilmesinde bir parametre seçimi gerekli iken geri yayılım algoritması için iki parametre, delta-bar-delta algoritması için üç parametre ve genişletil-

miş delta-bar-delta algoritması için dokuz parametre gerekli olmaktadır. Fazla sayıda parametre seçimi gerektiren algoritmalarda parametre değerlerinin yanlış belirlenme olasılığından kaynaklanan hatalar oluşmaktadır. Geri yayılım, delta-bar-delta, genişletilmiş delta-bar-delta algoritmalarında değer seçimi gerektiren parametre sayısının hızlı yayılım algoritmasına göre daha fazla olması, bu algoritmaların performanslarının düşük olmasının nedenlerindendir. Çok katmanlı perseptron sinir ağlarının eğitim ve test performans analizleri yapılmış ve sonuçlar değerlendirilmiştir. Hızlı yayılım algoritması ile eğitilen ağın minimum hataya yakınsaması diğer algoritmalar ile eğitilen ağlara göre daha az iterasyonda gerçekleşmiştir. Sınıflama sonuçları, alıcı işlem karakteristik eğrileri ve performans değerlendirme parametreleri, hızlı yayılım algoritması ile eğitilen çok katmanlı perseptron sinir ağının diyabet hastalığının teşhisinde en başarılı ağ olduğunu göstermiştir. Ağın bir defa eğitilmesinden ve testinden sonra ağın performansı dikkate alınarak otomatik teşhis yapılabileceği için performans değerlendirmesi önemlidir. Bu çalışmada sunulan ağ mimarisi ve ağ algoritmalarının performansı yeterli görüldüğü durumlarda hastaya ait ondört parametrenin ağ girişi olarak kullanılması ile sonuca ulaşılabilir.

KAYNAKLAR

1. Tafeit, E., Reibnegger, G., "Artificial Neural Networks in Laboratory Medicine and Medical Outcome Prediction", **Clinical Chemistry and Laboratory Medicine**, Vol 37, No 9, 845-853, 1999.
2. Güler, İ., Übeyli, E.D., "Detection of Ophthalmic Artery Stenosis by Least-Mean Squares Backpropagation Neural Network", **Computers in Biology and Medicine**, Vol 33, No 4, 333-343, 2003.
3. Übeyli, E.D., Güler, İ., "Neural Network Analysis of Internal Carotid Arterial Doppler Signals: Predictions of Stenosis and Occlusion", **Expert Systems with Applications**, Vol 25, No 1, 1-13, 2003.
4. Güler, İ., Übeyli, E.D., "ECG Beat Classifier Designed by Combined Neural Network Model", **Pattern Recognition**, Vol 38, No 2, 199-208, 2005.
5. Haykin, S., **Neural Networks: A Comprehensive Foundation**, New York, Macmillan College Publishing Company Inc, 1994.
6. Chaudhuri, B.B., Bhattacharya, U., "Efficient Training and Improved Performance of Multilayer Perceptron in Pattern Classification", **Neurocomputing**, Vol 34, 11-27, 2000.
7. Rumelhart, D.E., Hinton, G.E., Williams, R.J., "Learning Representations by Back-Propagating Errors", **Nature**, Vol 323, 533-536, 1986.
8. Jacobs, R.A., "Increased Rate of Convergence through Learning Rate Adaptation", **Neural Networks**, Vol 1, 295-307, 1988.
9. Minai, A.A., Williams, R.D., "Back-Propagation Heuristics: A Study of the Extended Delta-Bar-Delta Algorithm", **Proceedings of International Joint Conference on Neural Networks**, Vol 1, 595-600, San Diego, California, 17-21 June, 1990.
10. Fahlman, S.E., "An Empirical Study of Learning Speed in Backpropagation Networks", **Computer Science Technical Report**, CMU-CS-88-162, Carnegie Mellon University, Pittsburgh, 1988.
11. Chan, L-W., "Efficacy of Different Learning Algorithms of the Back Propagation Network", **IEEE Region 10 Conference on Computer and Communication Systems**, Vol 1, 23-27, 24-27 September, 1990.
12. Sidani, A., Sidani, T., "A Comprehensive Study of the Back Propagation Algorithm and Modifications", **IEEE Conference Record**, 80-84, 29-31 March, 1994.
13. Hannan, J.M., Bishop, J.M., "A Comparison of Fast Training Algorithms Over Two Real Problems", **IEE Fifth International Conference on Artificial Neural Networks**, Conference Publication No. 440, 1-6, 7-9 July, 1997.
14. Download Diabetes Data; 05.05.2005. <http://www.cormactech.com/neunet/download.html>.
15. Besser, G.M., Bodansky, H.J., Cudworth A.G., **Clinical Diabetes an Illustrated Text**, Gower Medical Publishing, London, 1988.
16. Kennedy, L., Lyons, T.J., "Non-Enzymatic Glycosylation", **British Medical Bulletin**, Vol 45, No 1, 174-190, 1989.
17. Bray, G.A., **Contemporary Diagnosis and Management of Obesity**, Newtown, Pennsylvania, Handsbook in Health Care Company, 2003.
18. Hyöty, H., Hiltunen, M., Lönnrot, M., "Enterovirus Infections and Insulin Dependent Diabetes Mellitus-Evidence for Causality", **Clinical and Diagnostic Virology**, Vol 9, No 2-3, 77-84, 1998.
19. Pi-Sunyer, F.X., "A Review of Long-Term Studies Evaluating the Efficacy of Weight Loss in Ameliorating Disorders Associated with Obesity", **Clinical Therapeutics**, Vol 18, No 6, 1006-1035, 1996.
20. American Diabetes Association (ADA), 05.05.2005 <http://www.diabetes.org/home.jsp>