Gazi Üniv. Müh. Mim. Fak. Der.
Cilt 30, No 1, 111-117, 2015

Journal of the Faculty of Engineering and Architecture of Gazi University
Vol 30, No 1, 111-117, 2015

İLERİ-GERİ TAKİP ALGORİTMASI TABANLI SEYREK AŞIRI ÖĞRENME MAKİNESİ

Ömer Faruk ALÇİN[*], Abdulkadir ŞENGÜR[**], Melih Cevdet İNCE[**]
*Fırat Üniversitesi, Teknik Eğitim Fakültesi, Elektronik ve Bilgisayar Eğitimi Bölümü, Elazığ
**Fırat Üniversitesi, Teknoloji Fakültesi, Yazılım Mühendisliği Bölümü, Elazığ
ofalcin@firat.edu.tr, ksengur@firat.edu. tr, mcince@firat.edu. tr

(Geliş/Received: 09.04.2014; Kabul/Accepted: 16.12.2014)

ÖZET

Son zamanlarda, Aşırı Öğrenme Makinaları (AÖM) makine öğrenmesi alanında ilgi duyulan bir konu haline gelmiştir. AÖM Tek Gizli Katmanlı İleri Beslemeli (TGKİB) ağlar için önerilmiş yeni bir öğrenme algoritmasıdır. AÖM modeli iyi genelleme performansı, aşırı hızlı öğrenme yeteneği ve düşük işlem karmaşıklığı gibi avantajlara sahiptir. Bu avantajlarının yanı sıra, AÖM bazı dezavantajlara sahiptir. Öncelikle, çıkış ağırlıklarının hesaplanmasında en küçük kareler minimizasyonu kullanıldığından aşırı örtüşme ile karşı karşıyadır. Bir diğer sakınca ise AÖM'nin doğruluğunun gizli katman nöron sayısına bağlı olmasıdır. Bu olumsuzluk geniş ölçekli problemlerde daha kötü bir hal almaktadır. Çalışmamızda, bu olumsuzlukları gidermek için seyrek tabanlı bir AÖM modeli önerildi. Seyrek tabanlı AÖM modelinde, çıkış ağırlıklarının seyrek temsilini hesaplamak için yinelemeli fırsatçı takip algoritması tabanlı İleri-Geri Takip (İGT) algoritması kullanılmıştır. İGT-AÖM olarak adlandırılan yöntem normal AÖM ile karşılaştırıldığında aşırı örtüşmeyen (ezberlemeyen) ağ yapısı, düşük hesap karmaşıklığı ve gizli katman nöronlarının kabul edilebilir bir sayıda olması gibi üstünlüklere sahiptir. Yaygın kullanılan sınıflandırma problemleri üzerine yapılan deneysel çalışmalar ışığında İGT-AÖM yöntemi dikkate değer üstünlükler göstermektedir. Ayrıca başarımın etkinliğini vurgulamak için çalışmanın devamında normal AÖM ve en küçük açısal regresyon (Least angle regression, LARS), en küçük mutlak daralma ve seçme operatörü (Least absolute shrinkage and selection operator, LASSO) ve Elastic Net gibi düzenlenmiş AÖM'ler ile önerilen metodun karşılaştırılması yapılmıştır.

Anahtar Kelimeler: AÖM, düzenlenmiş AÖM, seyreklik, ileri-geri takip, TGKİB.

FORWARD-BACKWARD PURSUIT BASED SPARSE EXTREME LEARNING MACHINE

ABSTRACT

Recently, the Extreme Learning Machine (ELM) becomes an interesting topic in machine learning area. The ELM has been proposed as a new learning algorithm for Single-Hidden Layer Feed forward Networks (SLFNs). The ELM structure has several advantageous such as good generalization performance, extremely fast learning ability and low computational process. Besides this advantageous, the ELM structure has some drawbacks. Firstly, the ELM encounters over-fitting problems because of using a least squares minimization in calculation of the output weights. The other drawback is about accuracy of the ELM. It depends on the number of hidden neurons. This situation is a big challenge in high-dimensional problems for some practical applications. In this paper, we propose a sparse ELM model to overcome the above-mentioned drawbacks. In the sparse ELM model, Forward-Backward Pursuit (FBP) algorithms based on greedy pursuit was used to obtain sparse representation of the output weights. The proposed method which is called FBP-ELM, has several benefits in comparing with the traditional ELM schemes such as avoiding over-fitting, low computational complexity and with adequate number of neurons in hidden layer. FBP-ELM shows its remarkable advantages when it is compared with the empirical studies on commonly used classification benchmarks. Moreover, a comparison with the original ELM and the other regularized ELM schemes such as Least-angle regression (LARS), Least absolute shrinkage and selection operator (LASSO) and Elastic Net is presented to show effectiveness of proposed FBP-ELM method.

Keywords: ELM, regularized ELM, sparsity, forward-backward pursuit, SLNFs.

1. GİRİŞ (INTRODUCTION)

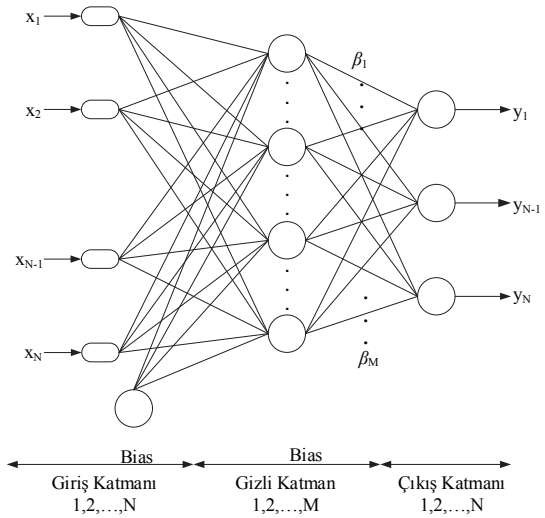
Son zamanlarda, Huang ve diğerleri tarafından önerilen AÖM örüntü tanıma, makine öğrenmesi ve veri madenciliği alanlarında bir hayli ilgi çekmiş ve konu ile ilgili çok sayıda uygulama gerçekleştirilmiştir [1-7]. AÖM, TGKİB ağları için yeni bir öğrenme algoritması olarak önerilmiştir [1]. Öğrenme sürecinde ağ ağırlıklarının gradyan temelli yenilenen TGKİB ağların aksine; AÖM'de, giriş ağırlıkları ve biaslar rastgele seçilir ve çıkış ağırlıkları ise analitik bir yöntemle hesaplanır. Bu durum AÖM'ye hızlı bir öğrenme, iyi bir genelleme performansı ve düşük bir hesaplama yükü gibi bazı önemli üstünlükler kazandırır. Bu önemli üstünlüklere rağmen, AÖM'e bazı dezavantajlara sahiptir [1-7]. İlk olarak, AÖM'nin çıkış ağırlıkları gizli katman çıkış matrisinin Moore-Penrose tersi bulunarak hesaplanır [1, 7, 8]. Bu durum, en küçük kareler minimizasyonu yönteminden dolayı aşırı örtüşmeye sebep olabilir [7, 9]. Ayrıca eğitim sürecinde kullanılan veri seti eğer tüm veri setinin özelliğini karakterize etmiyorsa bu problem çözülemez bir hal alır [7]. Diğer bir dezavantaj ise, AÖM'nin başarım performansının gizli katman nöron sayısına önemli ölçüde bağlı olmasıdır. Bahsedilen ilk dezavantajı gidermek için [7] ve [10] numaralı kaynaklarda, eğitim hatasını sınırlandıran L2-normu eklenerek, düzenlenmiş bir AÖM yöntemi önerilmiştir. Bu yöntem daha iyi genelleme performansı sağlarken, AÖM ağının daha karmaşık bir yapıda olmasına yol açar. Huang ve diğerleri [11] numaralı çalışmalarında yüksek genelleme performansına sahip bir çekirdek AÖM yöntemi önermişlerdir. Bu yöntem bahsedilen problemin çözümüne fayda sağlar gibi görünse de, büyük ölçekli uygulamalarda hesaplama yükü ve matris tersinin hesaplanması önemli bir problemi de beraberinde getirmiştir [12, 13]. Miche ve diğerleri tarafından önerilen yöntemde, en uygun ağ yapısının elde edilmesi için ilk önce gerekenden çok daha fazla sayıda gizli katman nöronu kullanılmış ve daha sonra belirli bir kriteri sağlamayan nöronlar budanarak yüksek genelleme performansına sahip optimum budamalı AÖM (OB-AÖM) yapısı oluşturulmuştur [7, 14]. OB-AÖM yönteminde fazla nöronların budanması işlemi L1-normlu regresyon yöntemi kullanılarak gerçekleştirilmiştir [7]. Miche ve diğerleri daha sonra OB-AÖM'ye L2-normu ekleyerek L1 ve L2-normlu kademeli yeni bir yöntem ile AÖM'nin başarım performansını iyileştirmişlerdir [4, 15]. Bu yeni yöntem Tikhonov'un düzenlenmiş optimum budamalı AÖM (TDOB-AÖM) olarak adlandırılmıştır [4, 15]. OB-AÖM ve TDOB-AÖM yöntemleri eğitim hatasının düşürülmesi ve genelleme performansının iyileştirilmesi gibi iyi özelliklere sahiptir. Fakat bu yöntemler ağın gizli katmanın çok fazla sayıda nörona sahip olmasına ve böylece hesaplama yükünün artmasına sebep olur [7]. Yukarıda bahsedilen düzenlenmiş AÖM yöntemleri, AÖM'nin bazı eksikliklerini giderse de optimum nöron sayısı ve hesaplama karmaşıklığı açısından bakıldığında hala yetersiz kalmaktadır.

Bu çalışmada, düşük hesaplama yükü ve en uygun gizli katman nöron sayısına sahip modeli elde etmek için seyrek tabanlı bir AÖM yöntemi önerilmiştir. Seyrek tabanlı AÖM yönteminde çıkış ağırlıklarının seyrek olduğu kabul edilir ve ağırlıklar seyrek geri çatma (SGÇ) yöntemi ile elde edilir. Diğer parametreler ise normal AÖM'de olduğu gibi rastgele üretilir. Böylece çıkış ağırlıklarının seyrek temsili elde edilerek uygun gizli katman nöron sayısına sahip model oluşturulur. SGÇ problemi için, L1-normlu eniyileme yöntemi, yinelemeli fırsatçı takip algoritmaları, Bayesian yöntemleri ve konveks olmayan optimizasyon yöntemi gibi çeşitli yaklaşımlar mevcuttur [16-18]. Bu çalışmada, SGÇ problemi için yinelemeli fırsatçı takip algoritmalarından İleri-Geri Takip (İGT) algoritması seçilmiştir. Önerilen seyrek tabanlı AÖM yöntemi İGT-AÖM olarak isimlendirilmiştir. Bu makalede, İGT-AÖM yöntemi biyo-medikal veri setlerinin sınıflandırılması problemlerine uygulanmıştır. Yukarıda bahsedildiği gibi, İGT-AÖM'nin genel amacı çıkış katmanı ağırlıklarının seyrek yaklaşım ile hesaplanarak düşük işlem karmaşıklığına ve en uygun gizli katman nöron yapısına sahip yapının elde edilmesidir. Öncelikle, gizli katman nöron yapısı AÖM'nin gerek duyduğundan daha büyük mimariye sahip ağ yapısı seçildi. Ardından, İGT algoritması yardımıyla çıkış katmanı ağırlıklarının seyrek temsili elde edilerek ilgisiz nöronlar kaldırılmıştır. Böylece, düşük işlem karmaşıklığına ve en uygun nöron sayısına sahip ağ yapısı elde edilmiştir. Makale, 5 bölüm halinde şu şekilde düzenlenmiştir: Bölüm 2'de Aşırı Öğrenme Makinaları ve İleri-Geri Takip algoritması hakkında teorik bilgiler verilmiştir. Önerilen İGT-AÖM yöntemi Bölüm 3'de detaylandırılmıştır. Bölüm 4'de Deneysel Çalışma sonuçları verilmiştir. Bölüm 5'te ise elde edilen sonuçlar irdelenmiştir.

2. TEORİK ALTYAPI (THEORETICAL BACKGROUND)

2.1. Aşırı Öğrenme Makinesi (Extreme Learning Machine)

Huang ve arkadaşları tarafından önerilen AÖM, TGKİB bir sinir ağı modeli için kullanılan bir öğrenme algoritmasıdır [1-3]. Gradyan temelli ileri beslemeli ağların aksine, AÖM'de giriş ağırlıkları ve eşik değerleri rastgele üretilirken, çıkış ağırlıklarının hesaplanmasında analitik yöntemler kullanılır [3]. Bu yöntem sayesinde öğrenme süreci hızlandırılmış olur. AÖM hızlı öğrenebilme yeteneğinin yanı sıra geleneksel geri yayılım algoritması ile öğrenen ileri beslemeli ağlara göre daha iyi genelleme başarımına sahiptir [1, 3]. Şekil 1'de TGKİB ağın örnek modeli gösterilmiştir. Şekil 1'de gösterilen TGKİB mimarisi için belirli bir girdi-çıktı ilişkisi x_i ve y_i ile gösterilsin.



Şekil 1. Tek gizli katmanlı ileri beslemeli ağ mimarisi (Single-Hidden layer feed forward network architecture)

Burada x_i , $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ matrisi, çıktı parametresi y_i ise $y_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T \in R^m$ matrisleri ile ifade edilir. *N* tane gizli sinir hücresi içeren ve *g(x)* aktivasyon fonksiyonuna sahip standart TGKİB, matematiksel olarak aşağıdaki gibi modellenir:

$$\sum_{i=1}^{M} \beta_i g\left(w_i . x_j + b_i\right) = o_j,\ j = 1, \dots, N \quad (1)$$

Burada $w_i = [\omega_{i1}, \omega_{i2}, \dots, \omega_{in}]^T$, *i*'ninci gizli sinir hücresine ve giriş sinir hücresine bağlı olan ağırlık vektörüdür, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ ise *i*'ninci gizli sinir hücresine ve çıkış sinir hücrelerine bağlı olan ağırlık vektörüdür. b_i ise *i*'ninci gizli sinir hücresinin eşik değeridir. $o_j = [o_1, o_2, \dots, o_j]^T$ise AÖM ağının çıkışıdır. $w_i . x_j$ ise w_i ve x_j'nin içsel çarpımını ifade eder. *N* tane gizli sinir hücresine ve *g(x)* aktivasyon fonksiyonuna sahip standart TGKİB yapısı ortalama "sıfır" hataya yaklaşabilir, $\sum_{j=1}^{M} \|o_j - y_j\| = 0$, β_i, w_i ve b_i arasındaki ilişki eşitlik (2)'de verilmiştir;

$$\sum_{i=1}^{M} \beta_i g\left(w_i . x_j + b_i\right) = y_j,\ j = 1, \dots, N \quad (2)$$

Denklem 2'de verilen *N* tane eşitlik, Denklem 3'deki gibi kısaltılabilir:

$$Y = H\beta \quad (3)$$

Burada;

$$H = \begin{bmatrix} g(w_1 . x_1 + b_1) & \cdots & g(w_N . x_1 + b_N) \\ \vdots & \cdots & \vdots \\ g(w_1 . x_M + b_1) & \cdots & g(w_N . x_M + b_N) \end{bmatrix}_{NxM} \quad (4)$$

Gizli katman çıkış matrisini ve

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_M^T \end{bmatrix}_{Mx1} \text{ ve } Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{Nx1} \quad (5)$$

sırasıyla çıkış katmanı ağırlıklarını ve çıkış vektörünü ifade eder [1, 19]. Bu durumda, w_i giriş katmanı ağırlıkları ve b_i gizli katman eşik değerleri rastgele seçilerek *H* analitik olarak hesaplanır. Geleneksel TGKİB ağlarda ağın eğitilmesi gerekirken AÖM'de eşitlik (3)'de belirtilen doğrusal bir denklem elde edilir. Çıkış ağırlıkları β;

$$\beta = H^{\dagger} Y \quad (6)$$

Burada β, eşitlik (6) çözdürülerek elde edilir. Burada $H^{\dagger}$, *H* çıkış matrisinin genelleştirilmiş tersi Moore-Penrose matrisi olarak tanımlanmıştır [1, 7, 8].

2.2. İleri-Geri Takip Algoritması (Forward-Backward Pursuit Algorithm)

Yinelemeli fırsatçı takip algoritmaları, SGÇ problemleri için alternatif bir yöntemdir [20].Yinelemeli algoritmalar, sinyali geri çatmak için tüm vektörleri kullanmak yerine uygun vektörleri kullandıkları ve bunların çarpım katsayılarının uygun seçiminden ötürü, işlem karmaşıklıkları az ve dolayısıyla hızlıdırlar [21-23].

Tablo 1. İleri-Geri takip algoritması (Forward-Backward pursuit algorithm)

Giriş H: Ölçüm Matrisi, *Y*: Ölçüm sinyali
Çıkış β: Çıkış ağırlıkları
Başlangıç $\boldsymbol{T^0 \leftarrow \emptyset}$, $\boldsymbol{r^0 \leftarrow Y}$, $\boldsymbol{k =\leftarrow 1}$
Döngü
İleri adım
$\boldsymbol{T_i = argmax_{J:\|J\|=\alpha} \left\|H_J^* r^{k-1}\right\|_1}$ İleri indeks setini oluştur.
$\boldsymbol{\tilde{T}^k = T^{k-1} \cup T_i}$ İndeks setini genişlet.
$\boldsymbol{w = \arg\min_w \left\|Y - H_{\tilde{T}^k} w\right\|_2}$
Geri adım
$\boldsymbol{T_g = arg\,min_{J:\|J\|=\lambda} \left\|w_J\right\|_1}$ Geri indeks seti oluştur.
$\boldsymbol{T^k = \tilde{T}^k - T_g}$ Geri indeks setlerini kaldır.
$\boldsymbol{w = \arg\min_w \left\|Y - H_{T^k} w\right\|_2}$
$\boldsymbol{r^k = Y - H_{T^k} w}$ rezidü güncelle.
Sonlandırma $\boldsymbol{\left\|r^k\right\|_2 \le \varepsilon \|Y\|_2}$ veya $\boldsymbol{\left|T^k\right| \ge K_{max}}$
Çıkış $\boldsymbol{\beta = w}$

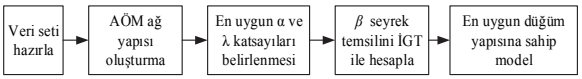
(ε: minimum hata toleransı, K_{max}: seyreklik seviyesi, T: indeks kümesi, r: rezidü, k: iterasyon sayısı)

İGT algoritması iki aşamalı eşiklemeli yinelemeli fırsatçı takip algoritmasıdır. İlk aşama olan ileri adımda, geri çatılmak istenen seyrek vektörün sıfır olmayan dizinlerini içeren destek kümesini ileri adım sayısı α kadar yeni dizin ekleyerek genişletir. İkinci aşama λ adım sayısından oluşan geri adımdır. Geri adımda destek kümesinden $\lambda<\alpha$ adet dizin çıkarılır. Böylelikle destek kümesi her döngüde genişletilmiş olur. İGT algoritması Tablo 1'de özetlenmiştir [16, 17]. İGT algoritması ile ilgili detaylı bilgi [16] çalışmasında verilmiştir.

Bu çalışmada İGT algoritması AÖM'nin en uygun gizli katman düğümlerini seçmek için kullanıldı. Böylece en uygun gizli katman düğümlü ağ yapısı şekillendirildi. Çıkış katmanı ağırlıklarının *N* boyutlu R^N düzleminde *K*-seyrek bir β sinyal olduğunu kabul edelim. Eşitlik (4)'e verildiği gibi *MxN* boyutlu gizli katman çıkış matrisi *H* hesaplayalım. *M* boyutlu bir *Y* ölçüm sinyalinin Eşitlik (3)'de belirtildiği gibi β vektörünün lineer ölçümü olduğunu varsayalım. Eşitlik (6)'nın İGT ile ters çözülmesi aradığımız çıkış katmanı ağırlıkları hesaplanacaktır.

3. İGT-AÖM YÖNTEMİ (FBP-ELM METHOD)

Bu bölümde, önerilen İGT-AÖM yöntemi kısaca anlatılmış ve deneysel çalışmalarla ilgili ayrıntılı bilgiler verilmiştir. İGT-AÖM yöntemi, blok şeması olarak Şekil 2'de gösterilmiştir.



Şekil 2. Önerilen İGT-AÖM yöntemi blok şeması (Block diagram of the proposed FBP-ELM method)

İGT-AÖM yönteminde, İGT algoritması durdurma kriteri olarak hata *(ε)* seçilmiştir. İGT algoritması için ileri adım sayısı α ve geri adım sayısı λ belirlenmesi önemlidir. Deneysel çalışmada en uygun α ve λ belirlenmesi için basit bir arama algoritması kullanılmıştır. Arama algoritmasında adım boyutlarının belirlenmesi [16] numaralı kaynakta gösterildiği gibi α, *M*'ye bağlı ve λ katsayısı ise α katsayısına bağımlı olarak seçilmiştir. Kullanılan arama algoritmasında α adım boyutu *0.1*M* ile *0.3*M* aralığında 1'er artımla taranmış ve λ adım boyutu ise *α-1* yapılmıştır. Arama algoritmasının her adımında İGT algoritması 100 kere çalıştırılarak en yüksek doğruluğa sahip α ve λ katsayıları seçilmiştir. Seçilen uygun adım parametreli İGT algoritması kullanılarak çıkış katmanı ağırlıklarının seyrek temsili hesaplanmıştır. Ardından oluşturulan İGT-AÖM yöntemi ile optimum model elde edilmiştir.

4. DENEYSEL ÇALIŞMALAR (EXPERIMENTAL STUDIES)

Bu bölümde, İGT-AÖM performansı UCI [24], Statlib [25] ve Kent Ridge Bio-medical Data Set Repository [26]'den alınan biyomedikal veri setleri kullanılarak test edilmiştir. Bu veri setlerinden yedi tanesi iki ve üç tanesi ise çok sınıflıdır. Tablo 2'te ilgili veri setlerinin özellikleri verilmiştir. [3, 5] çalışmalarında belirtildiği gibi, bu veri setleri örnekler ve değişkenlerin sayısı bakımından heterojen yapıya sahip olduklarından dolayı seçilmiştir. Deneysel çalışmada; doğruluk yüzdesi, standart sapma, kullanılan nöron sayısı ve deneysel çalışma süresi başarı ölçütü olarak ele alınmıştır.

Tablo 2. Kullanılan veri setlerinin özellikleri (Information about the datasets)

Veri Seti	Özellik Sayısı	Sınıf Sayısı	Örnekler	
			Eğitim	Test
Breast Cancer	10	2	615	68
Iris	4	3	135	15
Kidney	6	3	68	8
Liver	7	2	306	34
Lung Cancer	56	3	24	3
Lupus	3	2	78	9
Ovarian Cancer	15154	2	228	25
Parkinsons	23	2	175	20
Pima Indian Diabetes	8	2	691	77
Prostate Tumor	12600	2	91	10

Doğruluk yüzdesi, doğru tahmin edilen veri sayısının tüm veri sayısına oranının 100 katı ve Nöron Sayısı en yüksek doğruluk yüzdesine sahip nöron sayısı olarak tanımlanmıştır. Deneysel çalışmanın geçerliliğini artırmak için 10-Çapraz Geçerleme (10-fold crooss validation) tekniği kullanıldı ve elde edilen sonuçların ortalaması hesaplanmıştır. Modellerde kullanılan en yüksek doğruluk yüzdesine sahip en uygun nöron sayısını belirlemek için arama algoritması kullanıldı. Arama algoritmasında nöron sayısı beşten başlatılarak eğitim veri setinde kullanılan örnek sayısının onda bir oranında artış miktarı ile 1000 nörona kadar artırılarak en uygun nöron sayısı elde edilmiştir. Deneysel çalışmanın geçerliliğini sağlamak amacıyla her bir veri seti için benzetim 100 kere döndürüldü ve performans analizi için ortalama doğruluk yüzdesi, standart sapma, kullanılan nöron sayısı ve benzetim süresi hesaplandı. Veri setlerinin hepsi sıfır ortalama ve birim varyansa (zscore) sahip olacak şekilde normalize edildi [27]. Bütün AÖM'ler için gizli katman aktivasyon fonksiyonu sigmoid fonksiyonu, çıkış katmanı aktivasyon fonksiyonu ise doğrusal fonksiyon olarak seçildi. Deneyler 2.2 GHz işlemcili 8 GB belleğe sahip bir bilgisayarda gerçekleştirildi. İGT-AÖM sınıflandırıcı sonuçları Tablo 3'de verilmiştir.

Tablo 3. İGT-AÖM yöntemi sonuçları (Results of FBP-ELM method)

	İGT-AÖM			
Veri seti	Doğr. (%)	σ	Nöron Sayısı	Süre (s)
Breast C.	92,35	3,38	10	3,36
Iris	98	3,22	5	1,62
Kidney	44,28	22,78	49	1,02
Liver	70,58	9,3	10	1,81
Lung C.	95	14,81	17	1,91
Lupus	53,75	13,24	5	0,84
Ovarian C.	97,2	3,78	100	11,25
Parkinsons	90,00	6,77	13	2,45
Pima I. D.	74,736	7,22	10	3,96
Prostate T.	89,00	5,67	89	14,18

Tablo 3'te ortalama doğruluk yüzdesi Doğr. ile ve doğruluğun standart sapma ise σ ile gösterilmiştir. Tablo 3'ten görülebileceği üzere İGT-AÖM en iyi başarım sonucunu Iris veri seti için üretmiştir. Bunun yanı sıra Breast C., Ovarian C., Lung C. ve Parkinsons veri setleri için kabul edilebilir sonuçlar elde edilmiştir. Ayrıca en düşük standart sapmaya Iris veri setinde ulaşılmıştır. Kidney, Lung C. ve Lupus veri setleri için standart sapma değerleri en yüksek sonuçlar üretmiştir. Doğruluk ve standart sapma açısından en kötü sonuçlar ise Kidney veri seti için elde edilmiştir.

Deneysel çalışmalarda, İGT-AÖM başarım sonuçları analiz edildikten sonra yöntemin geçerliliğini göstermek için AÖM ve LARS-AÖM, LASSO-AÖM, Elastic Net-AÖM ile İGT-AÖM karşılaştırması yapılmıştır.

4.1. İGT-AÖM ile AÖM Karşılaştırılması (Comparison of FBP-ELM with ELM)

AÖM yönteminin başarımı da yine Tablo 2'deki veri setleri kullanılarak değerlendirilmiştir. Normal AÖM yöntemi ile İGT-AÖM yöntemi karşılaştırma sonuçları Tablo 4'de verilmiştir.

Tablo 4'den görüldüğü gibi genel olarak İGT-AÖM yöntemi AÖM yönteminden daha iyi sonuçlar vermiştir. Özellikle nöron sayısı bakımından İGT-AÖM yöntemi bütün veri setlerinde AÖM yöntemine üstünlük sağlamıştır. Breast C., Iris ve Pima I.D. veri setlerinde İGT-AÖM yöntemi tüm kriterlerde en iyi sonuçlara sahiptir.İGT-AÖ M yöntemi; Liver, Lupus ve Pima I.D. veri setleri için normal AÖM yönteminden yaklaşık iki katında daha yüksek doğruluk değerine sahiptir.

Ovarian C. ve Kidney veri setlerinde AÖM yöntemi İGT-AÖM yönteminden çok az bir fark ile daha iyi doğruluk yüzdesine ve standart sapmaya sahiptir. Lung C. Veri setinde her iki yöntem aynı doğruluk yüzdesi üretirken İGT-AÖM yöntemi yaklaşık olarak 17 kat daha az nöron sayısı ile bu sonucu elde etmiştir. Süre açısından sonuçlar incelendiğinde Lung C., Ovarian C., Parkinsons veri setleri dışında bütün veri setlerinde İGT-AÖM yöntemi iyi sonuçlar göstermektedir.

4.2. İGT-AÖM Yönteminin LARS, LASSO ve Elastic Net ile Karşılaştırılması (Comparison of FBP-ELM Method with LARS, LASSO and Elastic Net)

İGT-AÖM'nin başarımı ayrıca literatürde önerilen diğer LASSO, LARS, Elastic Net gibi düzenlenmiş AÖM yöntemleri ile karşılaştırılmıştır. Elde edilen ilgili sonuçlar Tablo 5'te gösterilmiştir.

Tablo 5'te görüldüğü gibi İGT- AÖM yöntemi Lupus ve Pima I. D. veri setleri dışındaki tüm veri setlerinde doğruluk açısından en iyi başarıma sahiptir. Ayrıca, Iris veri setinde İGT-AÖM yöntemi süre haricindeki

Tablo 4. İGT-AÖM ile AÖM karşılaştırması (Comparison of FBP-ELM with ELM)

	AÖM				İGT-AÖM			
Veri seti	**Doğr. (%)**	**σ**	**Nöron**	**Süre (s)**	**Doğr. (%)**	**σ**	**Nöron**	**Süre (s)**
Breast C.	74,11	5,46	917	8,01	**92,35**	**3,38**	**10**	**3,36**
Iris	72,66	9,13	785	9,36	**98**	**3,22**	**5**	**1,62**
Kidney	**45,71**	**14,75**	801	10,42	44,28	22,78	**49**	**1,02**
Liver	20,29	**7,77**	966	10,48	**70,58**	9,3	**10**	**1,81**
Lung C.	**95**	15,81	288	**0,44**	**95**	**14,81**	**17**	1,91
Lupus	17,5	16,87	72	1,75	**53,75**	**13,24**	**5**	**0,84**
Ovarian C.	**98,4**	**2,79**	900	**8,54**	97,2	3,78	**100**	11,25
Parkinsons	86,31	**5,08**	513	**1,02**	**90,00**	6,77	**13**	2,45
Pima I. D.	31,18	8,77	956	8,10	**74,736**	**7,22**	**10**	**3,96**
Prostate T.	**93**	8,23	990	**13,64**	89,00	**5,67**	**89**	14,18

Tablo 5. İGT-AÖM ile LARS, LASSO ve Elastic Net karşılaştırma sonuçları (Comparison of FBP-ELM Method with LARS, LASSO and Elastic Net)

	İGT-AÖM				LASSO				LARS				Elastic Net			
Veri Seti	**Doğr. (%)**	**σ**	**Nöron**	**Süre (s)**	**Doğr. (%)**	**σ**	**Nöron**	**Süre (s)**	**Doğr. (%)**	**σ**	**Nöron**	**Süre (s)**	**Doğr. (%)**	**σ**	**Nöron**	**Süre (s)**
Breast C.	**92,35**	3,38	**10**	3,36	82,94	6,83	673	1,57	88,82	3,33	795	**1,39**	90,29	**3,11**	917	59,36
Iris	**98**	**3,22**	**5**	1,62	96,66	4,71	405	**0,44**	92	6,88	405	1,28	96,66	3,51	270	3,97
Kidney	**44,28**	22,78	**49**	1,02	30	15,72	819	**0,66**	27,14	**10,54**	315	1,70	28,57	16,49	378	5,01
Liver	**70,58**	9,3	**10**	1,81	63,52	8,56	306	**0,38**	39,11	7,84	306	3,39	67,64	**6,79**	306	8,39
Lung C.	**95**	**14,81**	**17**	1,91	60	21,08	108	**1,9**	75	26,35	288	1,99	80	25,81	198	9,13
Lupus	53,75	**13,24**	**5**	0,84	**78,75**	14,49	864	**0,58**	56,25	13,51	576	1,14	63,75	17,12	216	4,52
Ovarian C.	**97,2**	**3,78**	**100**	11,25	90,4	5,39	675	**7,89**	86,8	9,62	675	10,41	93,2	3,79	900	14,44
Parkinsons	**90**	**6,77**	**13**	2,45	87,36	7,92	171	**0,41**	55,26	12,95	684	2,11	83,15	7,76	513	4,69
Pima I. D.	74,736	7,22	**10**	3,96	**76.31**	**6.73**	684	**1.19**	70,78	6,87	888	15,51	65	6,74	684	48,18
Prostate T.	**89**	**5,67**	**89**	14,18	76	8.43	540	**13.57**	72	19,88	810	15,06	78	11,07	450	17,07

bütün ölçütlerde en iyi sonuçları üretmiştir. İGT-AÖM yöntemi Kidney veri seti için diğer yöntemlerden yaklaşık bir buçuk kat daha iyi doğruluk değerine sahiptir. Yine Tablo 5'te görüldüğü gibi Elastic Net, Lupus ve Prostate T. veri setlerinde doğruluk ölçütünde kabul edilebilir sonuçlara sahiptir. Genel itibariyle LASSO deneysel çalışma süresi bakımından İGT-AÖM yönteminden daha iyi sonuç vermiştir. Bütün veri setleri için İGT-AÖM yöntemi en uygun nöron sayısına sahip yöntemdir. Tablo 5'den çıkarılacak genel bir yorumla İGT-AÖM yöntemi; doğruluk başarımı, standart sapma ve nöron sayısı bakımından LARS, LASSO ve Elastic Net yöntemlerinden daha iyi başarıma sahiptir.

5. SONUÇLAR (CONCLUSIONS)

Bu çalışmada, bir AÖM ağının çıkış ağırlıklılarının hesaplanması için SGÇ yöntemlerinden İGT algoritması kullanılmıştır. İGT algoritmasının önemli parametrelerinden olan α ve λ belirlenmesi işlemleri için en yüksek doğruluğu bulmayı amaçlayan bir arama algoritması kullanılmıştır. Ardından gizli katman nöron sayısı gerekenden fazla seçilerek ağ oluşturulmuştur. Önerilen İGT-AÖM yöntemi ile çıkış ağırlıklarının seyrek temsili elde edilerek ilgisiz nöronlar otomatik olarak budanmıştır. Deneysel çalışmalarda biyo-medikal veri setleri sınıflandırılması yapılmıştır. İGT-AÖM yönteminin üstünlüğünü göstermek için normal AÖM ve LARS, LASSO ve Elastic Net yöntemleri ile karşılaştırması yapılmıştır. Yapılan deneysel çalışmaların ışığında İGT-AÖM yöntemi ile amaçlandığı gibi yüksek doğruluk yüzdesine sahip en uygun nöronlu model elde edilmiştir. İGT-AÖM yöntemi sayesinde normal AÖM, LARS, LASSO ve Elastic Net yöntemlerinden dokuz kat daha az nöron sayısına sahip model oluşturulmuştur. İGT-AÖM yöntemi Standart sapma açısından %60 oranında normal AÖM, LARS, LASSO ve Elastic Net yöntemlerine üstünlük sağlamıştır. İGT-AÖM yönteminde kullanılan α ve λ katsayıları için kullanılan arama algoritmasından dolayı deneysel çalışma süresi açısından LASSO, LARS yöntemlerinin gerisinde kalırken, genel itibariyle normal AÖM ve Elastic Netten daha iyi sonuçlara sahiptir. Sonuç olarak, önerilen İGT-AÖM yöntemi genel itibariyle sınıflandırma doğruluğu en uygun nöron sayısı ve standart sapma kriterleri bakımından üstün bir başarıma sahipken deneysel çalışma süresi bakımından LARS ve LASSO yönteminin daha az başarıma sahiptir.

KAYNAKLAR (REFERENCES)

1. Huang, G.-B., Zhu, Q.-Y. ve Siew, C.-K., "Extreme learning machine: theory and applications", **Neurocomputing**, Cilt 70, No 1, 489-501, 2006.
2. Luo, M., and Zhang, K. "A hybrid approach combining extreme learning machine and sparse representation for image classification", **Engineering Applications of Artificial Intelligence**, Cilt 27, 228-235, 2014.
3. Ertugrul, O. F., Tagluk, M. E., Kaya, Y. ve Tekin, R. "EMG signal classification by extreme learning machine", **Signal Processing and Communications Applications Conference (SIU'13)**, Girne-KKTC, 24-24 Nisan 2013.
4. Miche, Y., Van Heeswijk, M., Bas, P., Simula, O. ve Lendasse, A., "TROP-ELM: a double-regularized ELM using LARS and Tikhonov regularization", **Neurocomputing**, Cilt 74, No 16, 2413-2421, 2011.
5. Deng, W., Zheng, Q. ve Chen, L., "Regularized Extreme Learning Machine", **Computational Intelligence and Data Mining (CIDM'09)**, Nashville, TN, USA, 389-395, March 30-April 2 2009.
6. Martínez-Martínez, J. M., Escandell-Montero, P., Soria-Olivas, E., Martín-Guerrero, J. D., Magdalena-Benedito, R. ve Gómez-Sanchis, J., "Regularized extreme learning machine for regression problems", **Neurocomputing**, Cilt 74, No 17, 3716-3721, 2011.
7. Luo, J., Vong, C. M. ve Wong, P. K., "Sparse Bayesian Extreme Learning Machine for Multi-classification", **Neural Networks and Learning Systems**, Cilt 25, No 4, 836-843, 2014.
8. Banerjee, K. S., "Generalized inverse of matrices and its applications", **Technometrics**, Cilt 15, No 1, 197-197, 1973.
9. Bishop, C. M., **Pattern recognition and machine learning**, Springer, New York, A.B.D., 2006.
10. Huang, G. B., Li, M. B., Chen, L. ve Siew, C. K., "Incremental extreme learning machine with fully complex hidden nodes", **Neurocomputing**, Cilt 71, No 4, 576-583, 2008.
11. Huang, G. -B., Zhou, H., Ding, X. ve Zhang, R., "Extreme learning machine for regression and multiclass classification", **Systems, Man, and Cybernetics, Part B: Cybernetics**, Cilt 42, No 2, 513-529, 2012.
12. Li, G., Wen, C., Li, Z. G., Zhang, A., Yang, F. ve Mao, K., "Model-Based Online Learning with Kernels", **Neural Networks and Learning Systems**, Cilt 24, No 3, 356-369, 2013.
13. Kivinen, J., Smola, A. J. ve Williamson, R. C., "Online learning with kernels", **Signal Processing**, Cilt 52, No 8, 2165-2176, 2004.
14. Miche, Y., Sorjamaa, A., Bas, P., Simula, O., Jutten, C. ve Lendasse, A., "OP-ELM: optimally pruned extreme learning machine", **Neural Networks**, Cilt 21, No 1, 158-162, 2010.
15. Alcin, O. F., Sengur, A., Qian, J. ve Ince, M. C., "OMP-ELM: Orthogonal Matching Pursuit-Based Extreme Learning Machine for Regression",**Journal of Intelligent Systems**, Cilt 24, No 1, 135-143, 2015.

16. Karahanoglu, N. B. ve Erdogan, H., "Compressed sensing signal recovery via forward–backward pursuit", **Digital Signal Processing**, Cilt 23, No 5, 1539-1548, 2013.

17. Karahanoglu, N. B. ve Erdogan, H., "Optimal forward-backward pursuit for the sparse signal recovery problem", **Signal Processing and Communications Applications Conference (SIU'13)**, Girne-KKTC, 24-26 Nisan 2013.

18. Tropp, J. A. ve Wright, S. J., "Computational methods for sparse solution of linear inverse problems", **Proceedings of the IEEE**, Cilt 98, No 6, 948-958, 2010.

19. Alcin, O. F., Sengur, A., Ghofrani, S. ve Ince, M. C., "GA-SELM: Greedy algorithms for sparse extreme learning machine", **Measurement**, Cilt 55, 126-132, 2014.

20. Needell, D., **Topics in Compressed Sensing**, PhD Dissertation, Univ. of California, Davis, Mathematics, 2009.

21. Karakus, C. ve Gurbuz, A. C., "Comparison of iterative sparse recovery algorithms", **Signal Processing and Communications Applications (SIU'11)**, Antalya-Türkiye, 20-22 Nisan 2011.

22. Temlyakov, V.N., "Greedy Algorithms and M-Term Approximation with Regard to Redundant Dictionaries", **Journal of Approximation Theory**, Cilt 98, No 1, 117–145 1999.

23. Tropp, J. A., "Greed is good: Algorithmic results for sparse approximation", **Information Theory**, Cilt 50, No 10, 2231-2242, 2004.

24. Frank, A. ve Asuncion, A., UCI machine learning repository, [http://archive.ics.uci. edu/ml], 2011.

25. Vlachos, P., StatLib Dataset Archive Carnegie Mellon University, [http://lib.stat.cmu.edu/ datasets], 2005

26. Li, J. ve Liu, H., Kent ridge bio-medical data set repository, Inst. for Infocomm Research. [http:// sdmc.lit.org.sg/GEDatasets/Datasets.html], 2002.

27. MATLAB version 7.10.0. Natick, Massachusetts: The MathWorks Inc., 2010.