



T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

METASEZGİSEL ALGORİTMALAR İLE AŞIRI ÖĞRENME MAKİNESİ PARAMETRELERİNİN OPTİMİZASYONU

Musa DOĞAN

YÜKSEK LİSANS

Bilgisayar Mühendisliği Anabilim Dalı

Temmuz-2020
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Musa DOĞAN tarafından hazırlanan "Metasezgisel Algoritmalar İle Aşırı Öğrenme Makinesi Parametrelerinin Optimizasyonu" adlı tez çalışması 21/07/2020 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Selçuk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri	**İmza**
Başkan Prof. Dr. Şakir TAŞDEMİR	……..
Danışman Dr. Öğr. Üyesi İlker Ali ÖZKAN	……..
Üye Dr. Öğr. Üyesi Onur İNAN	……..

Yukarıdaki sonucu onaylarım.

Prof. Dr. Sait GEZGİN
FBE Müdürü

ÖZET

YÜKSEK LİSANS TEZİ

METASEZGİSEL ALGORİTMALAR İLE AŞIRI ÖĞRENME MAKİNESİ PARAMETRELERİNİN OPTİMİZASYONU

Musa DOĞAN

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi İlker Ali ÖZKAN

2020, 75 Sayfa

Jüri
Dr. Öğr. Üyesi İlker Ali ÖZKAN
Prof. Dr. Şakir TAŞDEMİR
Dr. Öğr. Üyesi Onur İNAN

İleri beslemeli sinir ağları çok popüler ve tahmin içeren çoğu uygulamada sıklıkla kullanılan bir ağ mimarisidir. Bu mimarilerde nöronlar arası bilgi akışı girdi katmanından çıktı katmanına doğrudur. Aşırı Öğrenme Makinesi (AÖM), tek gizli katmana sahip ileri beslemeli bir yapay sinir ağı algoritmasıdır. AÖM'de geleneksel ileri beslemeli sinir ağlarından farklı olarak giriş ağırlıkları (w) ve gizli biaslar (b) rastgele belirlenmekte ve çıkış katmanındaki nöronların ağırlıkları analitik olarak Moore-Penrose adı verilen genelleştirilmiş matris tersimsisi yöntemiyle hesaplanmaktadır. Gradyan temelli algoritmalar ile sinir ağı parametrelerinin yinelemeli olarak ayarlanması yerine AÖM'de analitik olarak hesaplamayla öğrenme hızının ve genelleştirme performansının arttığı görülmüştür.

AÖM'de giriş ağırlığı ve gizli eşik değeri parametrelerinin rastgele seçilmesi çok hızlı bir şekilde öğrenme işleminin gerçekleşmesini sağlamakla beraber bazı durumlarda genelleştirme performansı iyi bir sonuç vermemektedir. Ayrıca AÖM, gradyan temelli algoritmalara göre gizli katmanda daha fazla nörona ihtiyaç duyabilmektedir. Bu nedenlerle giriş ağırlıkları ve gizli eşik değerlerinin optimum değerlerinin belirlenmesine yönelik çalışmalar mevcuttur. Bu çalışmada AÖM parametrelerinin optimum değerlerinin belirlenmesinde meta sezgisel optimizasyon algoritmaları kullanılmıştır. Belirlenen optimum giriş ağırlıkları ve gizli eşik değerleri kullanılarak AÖM'nin çıkış ağırlıkları analitik olarak hesaplanmaktadır.

AÖM'nin optimizasyonunda daha önce uygulanan Parçacık Sürü Optimizasyonu (PSO), Rekabetçi Sürü İyileştirici (RSİ) gibi sürü zeka algoritmalarına ek olarak Harris Şahinleri Optimizasyonu (HŞO) ile yeni bir AÖM optimizasyon modeli geliştirilmiştir. AÖM'nin optimizasyonu için geliştirilen bu hibrid modeller açık erişimli veri setleri üzerinde denenmiş ve elde edilen sonuçlar birbirleriyle karşılaştırılmıştır.

Anahtar Kelimeler: AÖM parametreleri optimizasyonu, aşırı öğrenme makinesi, harris şahinleri optimizasyonu, meta sezgisel algoritmalar, optimizasyon, sınıflandırma

ABSTRACT

MS THESIS

OPTIMIZATION OF EXTREME LEARNING MACHINE PARAMETERS WITH METAHEURISTIC ALGORITHMS

Musa DOĞAN

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF SELÇUK UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE IN COMPUTER ENGINEERING

Advisor: Asst. Prof. Dr. İlker Ali ÖZKAN

2020, 75 Pages

Jury
Asst. Prof. Dr. İlker Ali ÖZKAN
Prof. Dr Şakir TAŞDEMİR
Asst. Prof. Dr. Onur İNAN

Feedforward neural networks are a very popular and frequently used network architecture in many predictive applications. In these architectures, the information flow between neurons is from the input layer to the output layer. Extreme Learning Machine (ELM) is a feed forward neural network algorithm with a single hidden layer. Unlike conventional feedforward neural networks, the input weights (w) and weight bias (b) are randomly determined in ELM, and the weights of the neurons in the output layer are calculated analytically using the generalized matrix inversion method called Moore-Penrose. Instead of iteratively adjusting neural network parameters with gradient-based algorithms, it was observed that the learning speed and generalization performance increased with analytical computation in ELM.

Although random selection of input weight and hidden bias parameters in ELM enables the learning process to happen very quickly, generalization performance does not give a good result in some cases. In addition, ELM may require more neurons in the hidden layer compared to gradient-based algorithms. For these reasons, there are studies to determine the optimum values of input weights and hidden bias values. In this study, meta-heuristic optimization algorithms were used to determine the optimum values of the ELM parameters. The output weights of ELM are calculated analytically by using the determined optimum input weights and hidden bias values.

In addition to swarm intelligence algorithms such as Particle Swarm Optimization (PSO), Competitive Swarm Optimizer (CSO), which were previously applied in the optimization of ELM, a new ELM optimization model was developed with Harris' Hawks Optimization (HHO). These hybrid models developed for the optimization of ELM were tested on open access data sets and the results obtained were compared with each other.

Keywords: classification, ELM parameter optimization, extreme learning machine, harris' hawks optimization, meta-heuristic optimization, optimization

ÖNSÖZ

Lisans ve Yüksek Lisans eğitimim boyunca hoşgörüsünü ve desteğini eksik etmeyen, çalışmalarım sırasında beni yönlendiren değerli danışman hocam sayın Dr. Öğr. Üyesi İlker Ali ÖZKAN'a en içten teşekkürlerimi sunarım.

Manevi desteklerini eksik etmeyen kıymetli aileme, bilgi ve tecrübeleri ile bana destek olan bölümdeki tüm hocalarıma teşekkür ederim.

Musa DOĞAN
KONYA-2020

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vi**

İÇİNDEKİLER ..... **vii**

ŞEKİL LİSTESİ ..... **ix**

TABLO LİSTESİ ..... **xi**

SİMGELER VE KISALTMALAR ..... **xii**

1. GİRİŞ **1**

1.1. Literatür Araştırması 2
1.2. Tezin Amacı 5
1.3. Tez Organizasyonu 6

2. TEORİK ALT YAPI **8**

2.1. Yapay Sinir Ağları 8
2.2. Moore-Penrose Genelleştirilmiş Tersi 10
2.3. Minimum Norm EKK Çözümü 11
2.4. Aşırı Öğrenme Makinesi 11

3. OPTİMİZASYON ALGORİTMALARI **15**

3.1. Parçacık Sürü Optimizasyonu 15
3.2. Rekabetçi Sürü İyileştirici 16
3.3. Harris Şahinleri Optimizasyonu 18
3.3.1. Keşif Aşaması 19
3.3.2. Operasyon Aşaması 20

4. ÇALIŞMADA KULLANILAN VERİ SETLERİ **24**

4.1. Göğüs Kanseri Veri Seti 24
4.2. Diyabet Veri Seti 25
4.3. Şarap Veri Seti 26
4.4. İris Veri Seti 27
4.5. Tiroit Veri Seti 28

5. SINIFLANDIRMA İÇİN ÖNERİLEN HŞO-AÖM YÖNTEMİ **30**

5.1. Performans Değerlendirme Kriterleri 32
5.2. Uygunluk Fonksiyonu 34
5.3. Parametrelerin Belirlenmesi 34

6. DENEYSEL SONUÇLARIN İNCELENMESİ 36

6.1. İkili Sınıflandırma Problemlerinde Önerilen Model Başarımı 36
6.1.1. Göğüs Kanseri Veri Seti Sonuçları 36
6.1.2. Diyabet Veri Seti Sonuçları 41
6.2. Çoklu Sınıflandırma Problemlerinde Önerilen Model Başarımı 45
6.2.1. Şarap Veri Seti için Sonuçlar 45
6.2.2. İris Veri Seti için Sonuçlar 48
6.2.3. Tiroit Veri Seti için Sonuçlar 51

7. SONUÇLAR VE ÖNERİLER 54

KAYNAKLAR 57

ÖZGEÇMİŞ 61

ŞEKİL LİSTESİ

Şekil 2.1. 3 katmandan oluşan bir YSA modeli 8
Şekil 2.2. 4 katmandan oluşan ileri beslemeli bir YSA modeli 9
Şekil 2.3. 4 katmandan oluşan geri beslemeli bir YSA modeli 10
Şekil 2.4. Tek gizli katmanlı AÖM modeli 12
Şekil 2.5. AÖM adımları 13
Şekil 3.1. PSO'da parçacıkların hız ve konum güncellemeleri (Hassan ve ark 2005) 15
Şekil 3.2. RSİ'de parçacıkların güncelleme işlemleri (Cheng ve Jin 2015) 17
Şekil 3.3. HŞO algoritmasında keşif ve operasyon aşamaları grafiği (Heidari ve ark 2019) 18
Şekil 3.4. HŞO algoritmasının sözde kodu (Heidari ve ark 2019) 23
Şekil 4.1. Göğüs kanseri veri seti sınıf bazlı örnek sayıları 24
Şekil 4.2. Diyabet veri seti sınıf bazlı örnek sayıları 25
Şekil 4.3. Şarap veri seti sınıf bazlı örnek sayıları 26
Şekil 4.4. İris veri seti sınıf bazlı örnek sayıları 27
Şekil 4.5. Tiroit veri seti sınıf bazlı örnek sayıları 28
Şekil 5.1. HŞO-AÖM modeliyle parametrelerin optimizasyonu 31
Şekil 5.2. 5-katlamalı çapraz doğrulama yöntemi (Saporetti ve ark 2018) 33
Şekil 6.1. Modellerin göğüs kanseri veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi 37
Şekil 6.2. Modellerin göğüs kanseri veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi 37
Şekil 6.3. Göğüs kanseri veri seti için modellerin yakınsama grafikleri 40
Şekil 6.4. HŞO-AÖM modelinin göğüs kanseri veri seti üzerinde ROC eğrisi 40
Şekil 6.5. Modellerin diyabet veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi 41
Şekil 6.6. Modellerin diyabet veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi 41
Şekil 6.7. Diyabet veri seti için modellerin yakınsama grafikleri 44
Şekil 6.8. HŞO-AÖM modelinin diyabet veri seti üzerinde ROC eğrisi 45
Şekil 6.9. Modellerin şarap veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi 46

Şekil 6.10. Modellerin şarap veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi..... 46
Şekil 6.11. Modellerin iris veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi..... 49
Şekil 6.12. Modellerin iris veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi..... 49
Şekil 6.13. Modellerin tiroit veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi..... 51
Şekil 6.14. Modellerin tiroit veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi..... 52

TABLO LİSTESİ

Tablo 2.1. AÖM'de sıklıkla kullanılan aktivasyon fonksiyonları 11
Tablo 4.1. Göğüs kanseri veri setindeki özniteliklerin değer aralıkları ve ortalama değeri 25
Tablo 4.2 Diyabet veri setindeki özniteliklerin değer aralıkları ve ortalama değeri 26
Tablo 4.3. Şarap veri setindeki özniteliklerin değer aralıkları ve ortalama değeri 27
Tablo 4.4. İris veri setindeki özniteliklerin değer aralıkları ve ortalama değeri 28
Tablo 4.5. Tiroit veri setindeki özniteliklerin değer aralıkları ve ortalama değeri 29
Tablo 5.1. Çalışmada kullanılan veri setlerine ait özellikler 35
Tablo 5.2. Optimizasyon algoritmalarının parametreleri 35
Tablo 6.1. Modellerin göğüs kanseri veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri 38
Tablo 6.2. Göğüs kanseri veri seti için modellerin çalışmasından alınan sonuçlar 39
Tablo 6.3. Göğüs kanseri veri seti için modellerin hassasiyet, seçicilik, g-mean ve f-score sonuçları 39
Tablo 6.4. Modellerin diyabet veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri 42
Tablo 6.5. Diyabet veri seti için modellerin çalışmasından alınan sonuçlar 43
Tablo 6.6. Diyabet veri seti için modellerin hassasiyet, seçicilik, g-mean ve f-score sonuçları 44
Tablo 6.7. Modellerin şarap veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri 47
Tablo 6.8. Şarap veri seti için modellerin çalışmasından alınan sonuçlar 48
Tablo 6.9. Modellerin iris veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri 50
Tablo 6.10. İris veri seti için modellerin çalışmasından alınan sonuçlar 51
Tablo 6.11. Modellerin tiroit veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri 52
Tablo 6.12. Tiroit veri seti için modellerin çalışmasından alınan sonuçlar 53
Tablo 7.1. HŞO-AÖM modelinin literatürde bulunan diğer optimizasyon algoritmaları ile oluşturulan modeller ile karşılaştırılması 54

SİMGELER VE KISALTMALAR

Simgeler

b	:	Eşik değeri (bias)
$pbest$	:	PSO'da parçacığa ait en iyi çözüm
$gbest$	:	PSO'da tüm parçacıklara ait en iyi çözüm
c_1, c_2	:	PSO'da öğrenme katsayısı
v_i, x_i	:	Aday çözümler
φ	:	RSİ'de kontrol parametresi (phi)
β_i	:	AÖM'de çıktı ağırlığı
$(.)^{\dagger}$	:	Moore-Penrose işleci
w_i	:	AÖM'de giriş ağırlıkları
g	:	AÖM'de aktivasyon fonksiyonu
T	:	Maksimum iterasyon sayısı
t	:	Mevcut iterasyon
r	:	Kaçış başarımı
$\overline{x}$	:	Ortalama değer
m	:	RSİ'de parçacık sayısı
E	:	Kaçış enerjisi
LB	:	Alt sınır
UB	:	Üst sınır
J	:	HŞO'da tavşanın rastgele zıplama gücü

Kısaltmalar

AÖM	:	Aşırı Öğrenme Makinesi
CB	:	Korelasyon Tabanlı
CNN	:	Evrişimli Sinir Ağı
DE	:	Diferansiyel Evrim
DNO	:	Doğru Negatif Oranı
DPO	:	Doğru Pozitif Oranı
EKK	:	En Küçük Kareler
EOS	:	Topluluk Çevrimiçi Sıralı
ESA	:	Evrişimli Sinir Ağı
GA	:	Genetik Algoritma
GK	:	Guguk Kuşu Arama
GO	:	Genetically Optimized
GSO	:	Group Search Optimizer
GY	:	Geri Yayılımlı Sinir Ağları
HŞO	:	Harris Şahinleri Optimizasyonu
İGK	:	İyileştirilmiş Guguk Kuşu Arama
İGT	:	İleri Geri Takip Algoritması Tabanlı Seyrek
İPSO	:	İyileştirilmiş Parçacık Sürü Optimizasyonu
KGK	:	Kriptografik Gizlilik Koruyan
KKO	:	Karınca Koloni Optimizasyonu
KUE	:	Kendinden Uyarlamalı Evrimsel
LF	:	Levy Flight
LM	:	Levenberg-Marquardt
LOO	:	Birini Dışarıda Bırak
LU	:	Lower-Upper
MAE	:	Ortalama Mutlak Hata
MAPE	:	Ortalama Mutlak Yüzde Hata
MRSR	:	Çok Yanıtlı Seyrek Regresyon
MR	:	Manyetik Rezonans
OB	:	Optimum Budanmış
OS	:	Çevrimiçi Sıralı
PSO	:	Parçacık Sürü Optimizasyonu

RBF	:	Radyal Bazlı Fonksiyon
ReLU	:	Rektifiye Doğrusal Ünite
RGB	:	Red, green, blue
RMSE	:	Kök Ortalama Kare Hatası
ROC	:	İşlem Karakteristik Eğrisi
RSİ	:	Rekabetçi Sürü İyileştirici
SB	:	Seyrek Bayesci
TGKİB	:	Tek Gizli Katmanlı İleri Beslemeli Sinir Ağları
TSK	:	Takagi-Sugeno-Kang
YAB	:	Yerel Alıcı Bölgeleri
YSA	:	Yapay Sinir Ağı
YSE	:	Yinelemeli Sert Eşikleme

1. GİRİŞ

İleri beslemeli sinir ağları'nın yakınsama özelliği birçok araştırmacı tarafından incelenmiş ve evrensel yakınlaştırıcı oldukları kanıtlanmıştır (Hornik 1991, Ito 1992, Leshno ve ark 1993). Bu özellik, sinir ağlarını sınıflandırma ve regresyon problemlerinde uygulanabilir kılmıştır. Reel uygulamalarda sinir ağı sonsuz olmayan (sınırlı sayıda) bir eğitim verisiyle eğitilmektedir. Huang ve Babri (1998) yaptıkları çalışmada sınırlı sayıdaki eğitim verisinin fonksiyon yakınlaşması için gizli katman N adet nöron bulunan ve herhangi bir aktivasyon fonksiyonu kullanılan tek gizli katmanlı ileri beslemeli sinir ağlarının (TGKİB), N adet farklı örneği tam olarak öğrenebileceğini göstermişlerdir.

Girdi katmanını gizli katmana bağlayan giriş ağırlıkları ve gizli katman eşik değerlerinin hemen hemen tüm ileri beslemeli sinir ağlarının öğrenme algoritmalarında ayarlanması gerektiği unutulmamalıdır (Huang ve ark 2006). TGKİB'yi eğitmek için popüler öğrenme algoritmaları nispeten yavaştır, çünkü TGKİB'de tüm parametrelerinin yinelemeli olarak ayarlanması gerekir ve bu da hesaplama yükünü artırırken modelin yerel bir minimumda kolayca sıkışmasına neden olur. Bununla birlikte, çoğu sinir ağı uygulamasında gözlemlendiği gibi, ağların tüm parametrelerinin ayarlanması, öğrenmeyi karmaşık ve verimsiz hale getirebilir (Huang ve Chen 2007, Ding ve ark 2014, Sun ve ark 2017).

Huang ve ark (2006) tarafından TGKİB'nin verimliliğini artırmak için aşırı öğrenme makinesi (AÖM) adını verdikleri hızlı öğrenen bir sinir ağı algoritması geliştirilmiştir. Geleneksel sinir ağı öğrenme algoritmalarından farklı olarak, öğrenme oranı, öğrenme epok sayısı gibi kontrol parametrelerinin ayarlanmasına ihtiyaç duyulmaz. AÖM yerel bir noktada hatanın takılması, aşırı uyma gibi sorunlarla uğraşmadan doğrudan sonuca ilerler (Huang ve ark 2006, Deng ve ark 2009, Pouzols ve Lendasse 2010). Giriş ağırlıkları ve gizli katman eşik değerleri rastgele verilip çıktı ağırlıkları Moore-Penrose (MP) genelleştirilmiş tersi ile analitik olarak belirlenmektedir. AÖM yinelemeli olarak ayar yapmadan tamamen otomatik olarak uygulanır ve kullanıcının müdahalesi gerekmez (Ding ve ark 2014).

AÖM'nin öğrenme hızı diğer geleneksel metotlara göre çok yüksektir. Eğitim esnasında yineleme olmadan işlemlerin tek seferde yapılması çoğu uygulama için öğrenme aşamasının çok kısa bir süre içerisinde tamamlanmasını mümkün kılmaktadır

ve AÖM'nin gradyan bazlı öğrenme algoritmalarından genelleme performansı olarak daha iyi olduğu yapılan çalışmalarla gösterilmiştir (Huang ve ark 2006).

1.1. Literatür Araştırması

AÖM algoritması sunulduğu zamandan bu yana veri madenciliği, makine öğrenmesi, görüntü işleme ve diğer alanlarda geniş bir şekilde kullanılmıştır. (Huang ve ark 2015) Özellikle makine öğrenmesi alanında sınıflandırma ve regresyon problemlerinde AÖM'nin etkili bir algoritma olduğu yapılan çalışmalarla gösterilmiştir. (Huang ve ark 2011). Sağladığı avantajlardan dolayı farklı AÖM modelleri ve hibrid yapıları geliştirilmiştir.

Han ve ark (2013), yaptıkları çalışmada iyileştirilmiş parçacık sürü optimizasyonu (İPSO) algoritması ile, giriş ağırlıklarını ve gizli eşik değerlerini, doğrulama seti kullanılarak kök ortalama karelerinin bulunmasının yanı sıra çıktı ağırlıklarının normlarına göre optimize edilmesini sağlamıştır. Önerilen İPSO-AÖM algoritması, evrimsel aşırı öğrenme makinesi (E-AÖM), PSO-AÖM, AÖM ve Levenberg-Marquardt (LM) algoritmalarından daha iyi genelleme performansına sahiptir. Ayrıca önerilen algoritmanın E-AÖM, PSO-AÖM, AÖM ve LM algoritmalarından daha kararlı olduğu yapılan deneysel çalışmayla gösterilmiştir.

Matias ve ark (2013), TKGİB ağlar için genetik olarak optimize edilmiş AÖM (GO-AÖM) modelini sunmuşlardır. Bu modelde TGKİB'nin giriş değerleri, gizli katman konfigürasyonu, eşik değeri ve Tikhonov'un düzenlileştirme faktörü genetik algoritma (GA) ile optimize edilmiştir. Düzenlileştirme işlemi çözümü daha büyük normlardan uzak tutar ve TKGİB genelleştirme kapasitesinde iyileşme sağlayarak test verilerindeki performansı iyileştirir. Çıkış ağırlıkları ise tıpkı temel AÖM'de olduğu gibi en küçük kareler (EKK) ile elde edilmiştir. Modelin etkinliği ve performansı açık erişimli 5 veri seti üzerinde denenmiştir. Yapılan bu çalışmayla önerilen modelin gizli katmandaki nöron sayısının deneme yanılma yoluyla seçilmesine ihtiyaç kalmadığı ve giriş değişkenlerinin otomatik seçilerek ağın boyutu ve genelleştirme kapasitesinin iyileştirildiği gösterilmiştir.

Luo ve ark (2014), AÖM modelinin sınıflandırma problemlerindeki çıkış ağırlıklarının öğrenilmesinde seyrek bayesci yaklaşım sunmuşlardır. Seyrek bayesci AÖM (SBAÖM) adını verdikleri bu yöntemde ağın çıktılarını marjinal olasılığı hesaplanmakta ve öğrenme aşamasında gereksiz çoğu gizli nöronun otomatik olarak

budanmasıyla ağın başarımı artırılmaktadır. İlave olarak eğitim hatasının minimize edilmesi yerine marjinal olasılığın artırılması işlemiyle SBAÖM modeli eğitim verisinin gereğinden fazla fit edilmesinden (overfit) kaçınmış olur.

Yinelemeli Sert Eşikleme (YSE) algoritmasının en uygun gizli katman nöron sayısının elde edilmesi ve eğitim süreci karmaşıklığı gibi konularda AÖM ve düzenlenmiş AÖM yöntemlerine karşı üstünlüğü incelenmiştir. AÖM'de olduğu gibi, YSE-AÖM yönteminde de giriş katmanından gizli katmana bağlayan giriş ağırlıkları ve eşikler rastgele seçilmektedir. Daha sonra çıkış katmanı ağırlıkları YSE algoritması kullanılarak hesaplatılır ve böylece seyreklik sayesinde ilgisiz nöronlar otomatik olarak budanarak en uygun nöron sayısına sahip ağ elde edilmiş olur. Deneysel çalışmalardan elde edilen sonuçlar YSE tabanlı AÖM'nin, AÖM'ye üstünlüğünü göstermiştir (Alçin 2015).

Mohapatra ve ark (2015) çalışmalarında ikili sınıflandırma problemlerinden 4 medikal veri seti üzerinde önerdikleri İyileştirilmiş Guguk Kuşu Arama – AÖM (İGK-AÖM) modelini denemişlerdir. İGK modeli, Guguk Kuşu Arama (GK) algoritmasında sabit tutulan olasılık faktörü ve adım sayısının modifiye edilmesiyle oluşturulmuştur. Çalışmada sunulan ve karşılaştırılan modellerin performansı eğitim doğruluğu, test doğruluğu, karmaşıklık matrisi, İşlem Karakteristik Eğrisi (ROC), hassasiyet, seçicilik, G-mean ve F-score ölçütleriyle karşılaştırılmıştır. Önerilen modelin temel AÖM modeline göre daha iyi sonuç verdiği ve medikal veri sınıflandırma problemlerinde etkili olduğu gösterilmiştir.

Varol (2015) çalışmasında RGB videolarda hareket ve etkinlik tanıma, öznitelik çıkarmayı araştırmış ve öznitelikleri sınıflandırma işlemi için AÖM kullanmıştır. Gerçekçi ortamda büyük ölçekli hareket tanıma problemi için yeni bir yaklaşım önermektedir. Önerdiği bu yaklaşımda UCF101, THUMOS 2014 ve ChaLearn Looking at People 2014 Track veri kümelerinden ilkinde videolar sadece hareket içerecek şekilde kırpılmışken, diğer iki veri kümesindekiler zamansal olarak kırpılmamıştır. Çalışma sonucunda THUMOS 2014 veri kümesinde 102 hareket sınıfı üzerinde %63.37 oranında doğruluğa erişilmiştir.

Yeom ve Kwak (2017) yaptıkları çalışmada otomatik malumat ifadesi ile AÖM'yi kullanarak kısa vadede elektrik yükü tahmin sonucunu araştırmışlardır. if-then kurallarının oluşturulması için sistematik bir yaklaşım olması amacıyla Takagi-Sugeno-Kang (TSK) tabanlı bir AÖM kullanmışlardır. Bu yöntemde ilk olarak, bir rastgele bölümleme matrisi oluşturup, rastgele kümeleme için küme merkezlerini tahmin edilir.

Elde edilen küme merkezleri bulanık "if-then" kurallarının öncül parametrelerini belirlemek için kullanılır. Daha sonra, TSK bulanık tipinin doğrusal ağırlıkları, EKK yöntemi kullanılarak tahmin edilir. Bu doğrusal ağırlıklar TSK tasarımında ortaya çıkan parametreler olarak kullanılır. Tahmin yönteminin performansını ölçmek için, sırasıyla ortalama mutlak hata (MAE), ortalama mutlak yüzde hata (MAPE), r-kare yöntemi ve kök ortalama kare hatası (RMSE) gibi istatistiksel özellikler kullanılmıştır. Deneysel sonuçlarla, önerilen yöntemin, sigmoid, sinüs, radyal baz fonksiyonu ve rektifiye doğrusal ünite (ReLU) gibi dört aktivasyon fonksiyonuna sahip geleneksel bir AÖM ile karşılaştırıldığında iyi bir performans gösterdiğini ortaya koyulmuştur.

Çoğu uygulama alanı hem sınıflandırıcı modeli hem de veriler için güvenlik ve gizlilik gerektirmektedir. Çatak ve Mustacoglu (2018) gizli kalması gereken verilerin ifşa olmaması için AÖM ile gizlilik korumalı protokol yaklaşımı sunmuşlardır. Farklı uzunluklarda şifreleme/şifre çözme anahtarına sahip 4 farklı veri seti kullanarak önerilen protokollerin eğitim zamanı performans değerlendirilmesi test edilmiştir. Önerilen Kriptografik Gizlilik Koruyan AÖM (KGK-AÖM) modeliyle eğitim metodunda kısaca istemci giriş verisini şifreler, rastgele oluşturur ve bulut sistemine şifreli mesajı yükler. Bulut sistemi AÖM'nin en zaman alan kısmını, gizli katman çıkış matrisini hesaplayıp sonucu istemciye gönderir.

Ahila ve ark (2015), güç sistemi bozukluklarının tahmini ve sınıflandırmasında hızlı tahmin ve cevap yeteneğinden dolayı AÖM'nin araştırması yapılmıştır. Çalışmada önce wavelet dönüşümü kullanılarak özellik çıkarımı yapılmış ardından özellik alt kümesinin ve AÖM parametrelerinin seçilmesi sağlanarak çok sınıflı bir PSO-AÖM mimarisi önerilmiştir. Yapılan çalışmayla AÖM'nin başarımı, özellik uzayından seçilen en iyi ayırt edici özelliklerin otomatik olarak tespit edilmesi ve en uygun modelin seçilmesiyle artırılmıştır.

Duan ve ark (2018) Evrişimli Sinir Ağı (ESA) ve AÖM'nin avantajlarından yararlanmak için ESA-AÖM adını verdikleri bir yüz tanıma mimarisi önermişlerdir. Daha düşük zaman karmaşıklığında olan bu melez mimari, iki popüler veri seti kullanılarak yapılan deneyde, özdeş evrişimsel katmanlar içeren saf ESA'ya göre daha iyi bir performans elde etmiştir.

Eltaeib ve Mahmood (2018), beyin tümörü sınıflandırması ve tespiti için derin öğrenme tabanlı üç aşamalı bir yöntem önermişlerdir. Ön işleme, AÖM yerel alıcı bölgeleri (AÖM-YAB) tabanlı tümör sınıflandırması ve görüntü işleme tabanlı tümör bölgesi sınıflandırmasından oluşan bu üç aşamanın ikinci aşamasında kafatasından

alınan manyetik rezonans (MR) görüntüleri AÖM-YAB kullanılarak iyi huylu ve kötü huylu olarak sınıflandırılmıştır. Daha önceki çalışmalara göre önerilen metodun iyi bir performans gösterdiği ortaya koyulmuştur.

Eshtay ve ark (2018) tarafından giriş ağırlıkları ve gizli nöron eşik değerlerini optimize etmek için Rekabetçi Sürü İyileştirici (RSİ) algoritması kullanarak yeni bir model önerilmiştir. Önerilen bu modelle genelleme başarımının artırılması, sınıflandırıcının stabilize edilmesi ve az sayıda gizli nörona sahip bir ağ oluşturulmuştur. Bu model 15 tıbbi teşhis sınıflandırma probleminde denenmiş ve deneysel sonuçlara göre daha az sayıda gizli nöron ve daha yüksek stabilite ile daha iyi genelleme performansı sağlanabildiği gösterilmiştir.

Altan (2018) yaptığı çalışmada derin öğrenmenin analiz kapasitesi ile AÖM genelleme başarımı ve hızını birleştirmeyi gerçekleştirecek farklı ayrıştırma yöntemleriyle çalışabilen çok çekirdekli derin AÖM'yi önermiştir. KOAH teşhisi ve KOAH seviyelerinin belirlenmesi üzerine farklı modeller geliştirilmiş ve performansları hesaplanmıştır. Önerilen çok çekirdekli derin AÖM ile, LU ayrıştırma ve Hessenberg ayrıştırma tabanlı AÖM otomatik kodlayıcı çekirdeği gibi basit matematiksel çözümler ve tersine çevirme çözümleri kullanarak, derin öğrenmeden dikkate değer bir hızda daha iyi sınıflandırma performansına ve genelleme kapasitesine sahip olduğu yapılan bu çalışmayla gösterilmiştir.

Yapılan literatür araştırması analiz edildiğinde AÖM'nin makine öğrenmesi alanında diğer algoritmalara kıyasla verilerin daha hızlı eğitim sürecinden geçmesi ve genelleştirme performansının yüksek olması başlıca tercih edilme nedenlerinden olmuştur. Sunulduğu yıldan itibaren sürekli geliştirilen AÖM'nin gerek küçük ölçekli veri setleriyle gerekse orta ve büyük ölçekli veri setleriyle çalışmaları mevcuttur.

1.2. Tezin Amacı

Bu tez çalışmasının temel amacı, öğrenme algoritmalarından olan AÖM'de rastgele belirlenen giriş ağırlıklarının ve gizli eşik değerlerinin metasezgisel sürü zekâsı algoritmaları ve özellikle literatürde AÖM için henüz rastlanmamış Harris Şahinleri Optimizasyonu (HŞO) ile model oluşturularak optimal değerlerinin bulunmasıdır. HŞO, 2019 yılında sunulmuş, literatürde henüz yeni olan bir algoritmadır ve literatürde kabul görmüş diğer optimizasyon algoritmalarına kıyasla üstün sonuçlar sergilemesinden dolayı bu tez çalışmasının tercih sebebi olmuştur (Heidari ve ark 2019).

Ayrıca AÖM için literatürde daha önce denenmiş ve iyi sonuçlar verdiği belirlenmiş, metasezgisel algoritmalardan Parçacık Sürü Optimizasyonu (PSO) ve Rekabetçi Sürü İyileştirici (RSİ) tabanlı AÖM modelleriyle karşılaştırılmıştır. Modelin eğitilmesi ve test edilmesinde sınıflandırma problemlerinde yaygın olarak kullanılan açık erişimli beş farklı veri seti üzerinde modellerin başarımın ölçülmesi için denemeler yapılmış analizler gerçekleştirilmiştir.

1.3. Tez Organizasyonu

Bu yüksek lisans tezi toplamda yedi ana bölümden oluşmaktadır.

Birinci bölümde giriş bölümü olarak tez konusunun genel tanımı ve motivasyonu verilmiş, konu ile literatürde benzer çalışmalar hakkında araştırmalar yapılmış ve kısaca analizi yapılmıştır.

İkinci bölümde konu ile ilgili genel bilgilerin verildiği birinci bölüme takiben sınıflandırma için kullanılan yapay sinir ağlarının tanıtımı, Moore-Penrose Genelleştirilmiş Tersi, Minimum Norm EKK çözümü ve bu tez çalışmasında sınıflandırıcı olarak kullanılan aşırı öğrenme makinesinin modelinin üzerinde durulmuştur. Modelde sıklıkla kullanılan aktivasyon fonksiyonları ve modelin sunulduğu günden bu yana geliştirilen farklı yaklaşımlarından bahsedilmiştir.

Üçüncü bölümde Optimizasyon Algoritmaları başlığı altında sürü zekası temelli Parçacık Sürü Optimizasyonu (PSO) ve Rekabetçi Sürü İyileştirici (RSİ) algoritmalarının çalışması hakkında bilgi, alanda yapılan çalışmalar ve tezin asıl konusu olan Harris Şahinleri Optimizasyonu (HŞO) algoritmasının genel yapısıyla beraber çalışma mantığı, keşif ve operasyon stratejileri verilmiştir.

Dördüncü bölümde çalışmada sınıflandırma için kullanılan açık erişimli beş farklı veri setinin (diyabet, iris, tiroit, göğüs kanseri ve şarap) özellikleri detaylı bir şekilde incelenmiştir.

Beşinci bölümde sınıflandırıcı için önerilen HŞO-AÖM modelinin eğitilmesinde kullanılan parametreler, aktivasyon fonksiyonları ve uygunluk değerinin seçilmesi gibi konular ele alınmıştır.

Altıncı bölümde önerilen HŞO-AÖM modelinin beşinci bölümde verilen veri setleri ile sınıflandırması yapılarak performansı analiz edilmiş, PSO-AÖM, RSİ-AÖM ve temel AÖM modeliyle karşılaştırılması yapılmış, literatürdeki yapılmış çalışmalarla karşılaştırılmış, sonuca etki eden faktörler incelenmiştir.

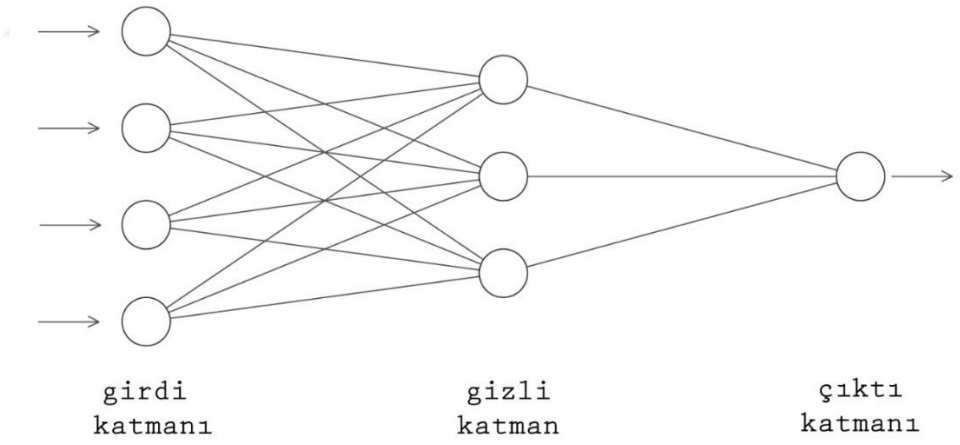
Yedinci Bölümde yapılan bu tez çalışmasıyla ortaya koyulan sonuçlar yorumlanmış, bu konuda çalışmak isteyen araştırmacılara önerilerde bulunulmuştur.

2. TEORİK ALT YAPI

Tezin giriş kısmında da belirtildiği üzere AÖM giriş ağırlıkları ve gizli eşik değerleri rastgele olarak verilmiş tek gizli katmana sahip ileri beslemeli bir YSA modelidir. Bu sebeple AÖM'nin algoritmasından bahsetmeden önce YSA hakkında bilgi verilmiştir.

2.1. Yapay Sinir Ağları

Yapay Sinir Ağları (YSA), makine öğrenmesi algoritmalarından biridir. İnsanların öğrenme şeklini taklit etmek için beynin çalışma mantığından esinlenilerek oluşturulmuştur (Zhang ve ark 1998). Bir sinir ağının davranışı, nöronlarının transfer fonksiyonları, öğrenme kuralı ve mimarinin kendisi tarafından belirlenir (Agatonovic-Kustrin ve Beresford 2000). YSA özellikle örüntü tanıma ve sınıflandırma algoritmalarında kullanılan yapay zeka yöntemlerindendir. Şekil 2.1.'de bir YSA örneği verilmiştir.

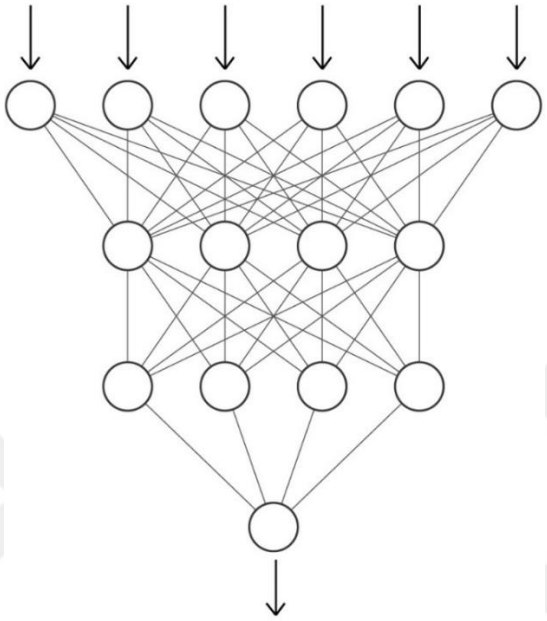


Şekil 2.1. 3 katmandan oluşan bir YSA modeli

Sinir ağlarının birçok çeşidi olmasına rağmen temel kuralları birbirine çok yakındır. Ağdaki her nöron giriş sinyalini alır, işlem yapar ve çıktı sinyali olarak gönderir. Bu algoritmada nöronlar arası bağlantıların önem derecesi ağırlık katsayısı olarak isimlendirilir (Svozil ve ark 1997).

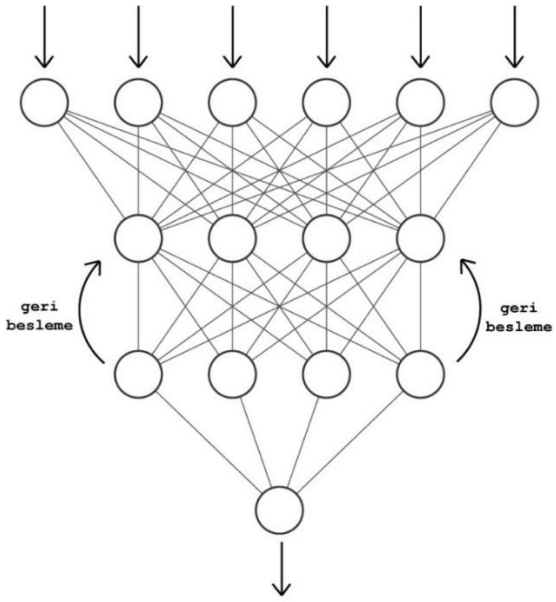
Nöronların kendi aralarındaki bağlantılarına göre YSA'lar ileri beslemeli ve geri beslemeli olarak ikiye ayrılmaktadır. İleri beslemeli ağlar çok popüler ve tahmin içeren

çoğu uygulamada sıklıkla kullanılan bir ağ yapısıdır. Bu mimarilerde nöronlar arası bilgi akışı çıktı katmanından girdi katmanına doğru ilerlemez. Dolayısıyla önceki nöronların çıktı kayıtlarını saklamaz. İleri beslemeli bir ağın yapısı Şekil 2.2.'de gösterilmiştir.



Şekil 2.2. 4 katmandan oluşan ileri beslemeli bir YSA modeli

Geri beslemeli ağlarda ise çıktı katmanından girdi katmanına doğru nöronlar arası bilgi akışı gerçekleşir. Bu yapıda bir ağda önceki durumun kaydı tutulacağı için çıktıya sadece giriş sinyalleri değil, aynı zamanda ağın bir önceki durumu da ilave olarak yansır (Agatonovic-Kustrin ve Beresford 2000). Geri beslemeli ağlara ait bir örnekte Şekil 2.3.'te verilmiştir.



Şekil 2.3. 4 katmandan oluşan geri beslemeli bir YSA modeli

2.2. Moore-Penrose Genelleştirilmiş Tersi

Lineer cebirde bir A matrisinin Moore-Penrose tersi ($A^{\dagger}$), ters matris kavramının genelleştirme olarak bilinen en yaygın türlerinden biridir. (Moore 1920, Bjerhammar 1951, Penrose 1955) bir diğer adı pseudoinverse olarakta adlandırılan genelleştirilmiş ters matrisini bağımsız olarak literatürde ilk kez kullanan isimlerdir. Bu matrisin tanımı matlab'da $pinv$ fonksiyonu şeklinde yerini almıştır. Bir matrisin genelleştirilmiş ters matris olabilmesi için Denklem (2.1)'deki virgülle ayrılmış bütün koşulları sağlaması gerekmektedir (Penrose 1955). Matrisin devriği (transpozu) $(.)^{*}$ ile gösterilmiştir.

$$AA^{\dagger}A = A,\ A^{\dagger}AA^{\dagger} = A^{\dagger},\ (AA^{\dagger})^{*} = AA^{\dagger}\ ,\ (A^{\dagger}A)^{*} = A^{\dagger}A \tag{2.1}$$

Pseudoinverse'in en yaygın kullanımı lineer denklem sistemlerinin çözümünde en küçük kareler (EKK) yönteminin en iyi çözümünü bulmaktır (Penrose 1956). A matrisi, tekil hatta kare olmayan bir matris olsa bile Moore-Penrose ile kolayca çözülebilmektedir (Huang ve ark 2006).

2.3. Minimum Norm EKK Çözümü

$Ax = y$ lineer denklemi için, EKK çözümü olan $\hat{x}$, Denklem (2.2)'deki gibi ifade edilebilir. Bu denklemde $\|\cdot\|$ Öklid uzayında normu belirtir (Huang ve ark 2006).

$$\| A\hat{x} - y \| = \min_{x} \| Ax - y \| \tag{2.2}$$

Reel sayılarda tanımlı herhangi bir y değeri için $Ax = y$ lineer sisteminin tüm EKK çözümleri arasındaki minimum norm EKK çözümü $x_0 = A^{\dagger}y$ şeklinde ifade edilir (Huang ve ark 2006). Bu tanım Denklem (2.3)'te gösterilmiştir.

$$\begin{aligned} &\| x_0 \| \leqslant \| x \|, \\ \forall x \in \{x : &\| Ax - y \| \leqslant \| Az - y \|, \forall z \in R^n\} \end{aligned} \tag{2.3}$$

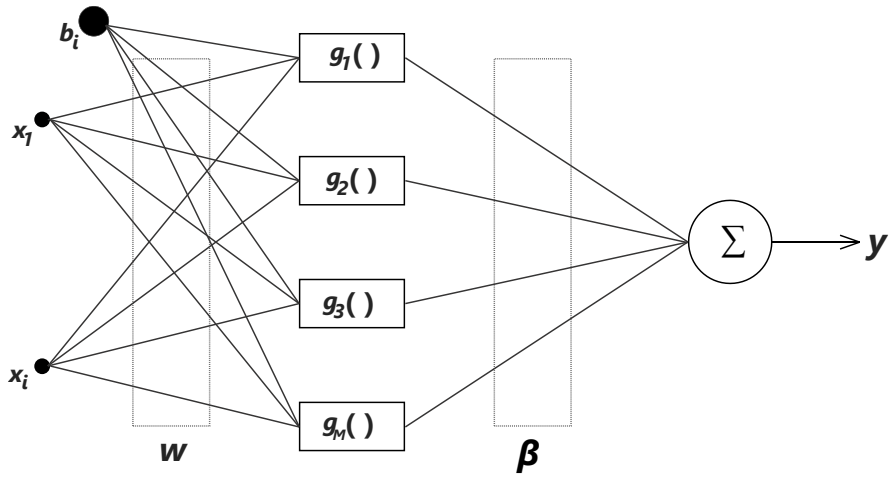
2.4. Aşırı Öğrenme Makinesi

Aşırı Öğrenme Makinesi (AÖM), hem regresyon hem de sınıflandırma uygulamalarında performansıyla dikkat çeken tek gizli katmana sahip ileri beslemeli öğrenme algoritmasıdır. Gizli katmandaki nöronlar için Sigmoid, Sin, Gaussian gibi aktivasyon fonksiyonları, çıktı katmanındaki nöronlar içinde lineer bir aktivasyon fonksiyonu kullanılır. En sık kullanılan aktivasyon fonksiyonları Tablo 2.1.'de verilmiştir.

Tablo 2.1. AÖM'de sıklıkla kullanılan aktivasyon fonksiyonları

Sigmoid Fonksiyon	$g(w,b,x) = \dfrac{1}{1 + exp\,(-(w \cdot x + b))}$
Hiperbolik Tanjant Fonksiyonu	$g(w,b,x) = \dfrac{1 - \exp\,(-(w \cdot x + b))}{1 + \exp\,(-(w \cdot x + b))}$
Gauss Fonksiyonu	$g(w,b,x) = \exp\,(-b \| x - w \|)$
Çoklu Kuadratik Fonksiyon	$g(w,b,x) = (\| x - w \| + b^2)^{1/2}$
Hard Limit Fonksiyonu	$g(w,b,x) = \begin{cases} 1, & if\ w \cdot x + b \leq 0 \\ 0, & \text{otherwise} \end{cases}$
Kosinüs Fonksiyonu	$g(w,b,x) = \cos\,(w \cdot x + b)$

Tek gizli katmana sahip AÖM modeli Şekil 2.4.'te gösterilmiştir.



Şekil 2.4. Tek gizli katmanlı AÖM modeli

Standart ileri beslemeli bir sinir ağında gizli katmanda M tane nöron ve N tane örnek bulunan bir modelin matematiksel gösterimi Denklem (2.4)'te verilmiştir.

$$\sum_{i=1}^{M} \beta_i g\left(w_i \cdot x_j + b_i\right), j = 1, \dots, N. \tag{2.4}$$

burada β_i çıkış ağırlıkları, g aktivasyon fonksiyonu, w_i giriş ağırlıkları ve b_i 'de eşik değeridir. Yukarıdaki eşitlikte gizli katman çıkış matrisi $H = g\left(w_i \cdot x_j + b_i\right)$ olarak ifade edilirse sade bir şekilde Denklem (2.5) elde edilir.

$$H\beta = y \tag{2.5}$$

H matrisi Denklem (2.6)'da, β ve y değerleri Denklem (2.7)'de verilmiştir.

$$\begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_M \cdot x_1 + b_M) \\ \vdots & \cdots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_M \cdot x_N + b_M) \end{bmatrix}_{NxM}, \tag{2.6}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_M \end{bmatrix}_{Mx1} \quad ve \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}_{Nx1} \tag{2.7}$$

H matrisi'nin $i.$ sütunu, $x_1, x_2, \dots, x_N$ girişlerine göre $i.$ gizli nöron çıkışıdır. AÖM yaklaşımında $w_i \; ve \; b_i$ değerleri rastgele atanır ve çıkış ağırlıkları Denklem (2.8)'de gösterilen denklem ile hesaplanır.

$$\hat{\beta} = H^{\dagger} y \tag{2.8}$$

Denklem (2.8)'de verilen denklemde $\hat{\beta}$ yaklaşık çıkış ağırlığı ve $H^{\dagger}$, H matrisinin genelleştirilmiş ters Moore-Penrose matrisi olur. Denklem (2.8)'i ifade etmenin bir başka yolu ise Denklem (2.9)'daki gibidir.

$$||H\hat{\beta} - \; y = 0|| \tag{2.9}$$

Temel AÖM algoritması Şekil 2.5.'te verilen 3 adımda özetlenebilir:

Aşırı Öğrenme Makinesi Adımları
1. **Adım:** Gizli katman nöron sayısını tanımlama, giriş parametrelerine (ağırlık $\boldsymbol{w_i}$ ve eşik değeri $\boldsymbol{b_i}$) rastgele atama yapma
2. **Adım:** $\boldsymbol{H}$ gizli katman çıktı matrisini hesaplama
3. **Adım:** Denklem (2.8)'e göre $\widehat{\boldsymbol{\beta}}$ çıktı ağırlığını hesaplama

Şekil 2.5. AÖM adımları

AÖM'nin sunulduğu yıldan bu yana pek çok geliştirilmiş versiyonu bulunmaktadır. Huang ve ark (2006) sundukları Artımlı AÖM (A-AÖM) yöntemi ile yeni bir gizli düğüm eklenmesi sırasında mevcut gizli düğümlerle olan çıkış ağırlıklarının dondurulmasıyla TGKİB'lerin giriş ağırlıklarıyla gizli katman eşik değerinin ve Radyal Bazlı Fonksiyon (RBF) ağlarının RBF düğümlerinin etki faktörlerinin ayarlanmasına gerek kalmadığını yaptıkları çalışmayla ortaya koymuşlardır.

AÖM giriş ağırlıklarını ve gizli eşik değerlerini rastgele seçtiğinden, geleneksel TGKİB ağlarında parametrelerin ayarlanması için harcanan öğrenme zamanından bir hayli tasarruf edilir fakat çıktı ağırlıkları, önceden belirlenmiş girdi ağırlıklarına ve gizli eşik değerlerine dayanarak hesaplandığı için, bir dizi istenmeyen veya gereksiz girdi ağırlıkları ve gizli eşik değerlerinin olması performansa kötü etki eder. Ayrıca AÖM, bazı uygulamalarda, bilinmeyen test verilerine yavaş yanıt veren geleneksel ayar tabanlı öğrenme algoritmalarından daha fazla gizli nöronun bulunduğu daha kompakt ağlar

gerektirmektedir. Zhu ve ark (2005), yaptıkları çalışmada kompakt ağlardaki başarımı artırmak için diferansiyel evrim (DE) algoritmasının modifiye edilmiş formunu kullanarak optimal giriş ağırlıkları ve gizli eşik değerlerini aramışlardır. Aldıkları sonuçlarla Ghosh ve Verma (2003) tarafından önerilen GALS metodu, temel AÖM ve geri yayılımlı (GY) sinir ağları'na göre daha iyi bir genelleştirme performansı elde edilmiştir. (Zhu ve ark 2005).

Birçok gerçek dünya uygulamasında verilerin tamamı aynı anda mevcut olmayabilir ve öğrenme işleminin de bir yandan gerçekleşmesi gerekebilir. Toplu öğrenme'de yeni veriler eski verileride kullanarak yeniden eğitim gerçekleştirir ve buda çok zaman almaktadır. Liang ve ark (2006) yaptıkları çalışmada eğitim verisinin sadece tek tek değil yığınlar halinde işlenebildiği çevrimiçi sıralı-AÖM (OS-AÖM) modelini sunmuşlardır. Lan ve ark (2009), OS-AÖM modelinin performansını artırmak ve topluluk ağlarında kullanılması için topluluk çevrimiçi sıralı-AÖM (EOS-AÖM) olarak adlandırılan entegre bir ağ yapısı önermişlerdir. Topluluktaki her bir OS-AÖM modelinin çıkış ortalama değeri, ağ performansının nihai ölçümü olarak kullanılmaktadır. EOS-AÖM modeli, tekli OS-AÖM modeli ile kıyaslandığında daha istikrarlı ve daha doğru sonuçlar vermiştir.

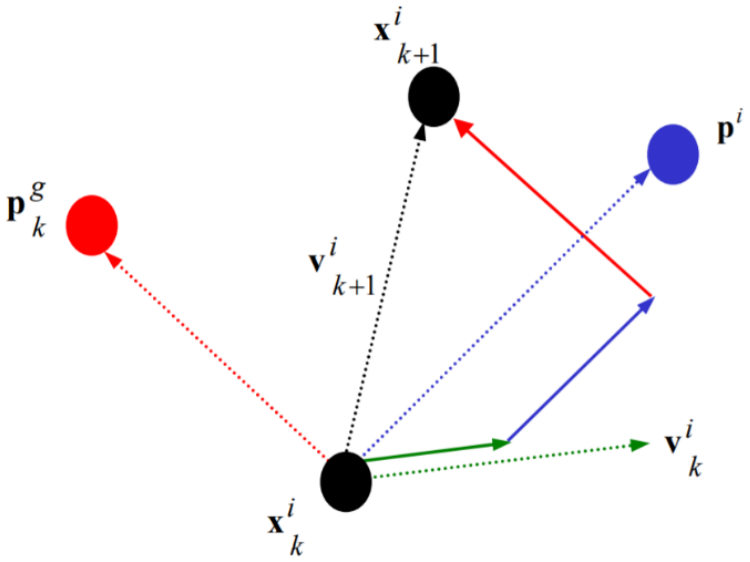
Birbirinden alakasız veya ilişkili giriş değişkenleri olduğunda AÖM bazı sorunlarla karşılaşabilir. Bu sorunu gidermek için optimum budanmış AÖM (OB-AÖM) metodolojisi önerilmiştir. OB-AÖM modeli temel AÖM algoritmasını genişletir ve nöronların budanması yöntemiyle daha sağlam bir algoritma ortaya çıkar. Üç adımdan oluşan bu algoritmada önce temel AÖM ile model oluşturulur ardından çok yanıtlı seyrek regresyon (MRSR) ile gizli katmanda bulunan nöronlar sıralamaya sokularak kullanışsız nöronlar elenir ve son olarak birini dışarıda bırak (LOO) yöntemi ile optimal nöron seçilir (Miche ve ark 2010).

3. OPTİMİZASYON ALGORİTMALARI

Bu bölümde çalışmada kullanılan sürü zekâsı algoritmalardan Parçacık Sürü Optimizasyonu, Rekabetçi Sürü İyileştirici ve Harris Şahinleri Optimizasyonu algoritmaları açıklanmıştır.

3.1. Parçacık Sürü Optimizasyonu

PSO, Kennedy ve Eberhart (1995) tarafından sunulan sürü zekasına dayalı popülasyon tabanlı optimizasyon algoritmasıdır. Doğada sürü halinde bulunan ve yine sürü halinde hareket eden karınca, kuş, arı ve balık gibi hayvanların sosyal ve fiziksel davranışlarından esinlenilerek geliştirilmiştir. Sürüdeki her bir bireyin parçacık olarak nitelendirildiği ve her bir parçacığın tek başına gerçekleştiremeyeceği karmaşık görevleri bütün parçacıklarla oluşan sürünün topluca başarabildiği akıllı bir sistemdir. PSO'da parçacıkların hız ve konum güncellemeleri Şekil 3.1'de gösterilmiştir.



Şekil 3.1. PSO'da parçacıkların hız ve konum güncellemeleri (Hassan ve ark 2005)

Popülasyon tabanlı olan bu tür algoritmalar birçok parçacıktan oluşur. Her bir parçacık için hız ve konum bilgileri tutulur. Vektörel olarak tutulan bu bilgiler optimizasyon problemlerinde: Hız vektörü v pozisyon değiştirme hızını yansıtırken, konum vektörü x parçacığın rastgele üretilen konum değerini niteler. Örneğin bir kuş parçacığı için hız vektörü bilgisi kuşun yer değiştirme miktarını belirtirken konum

bilgisi kuşun uçtuğu noktayı belirtir. Bu bilgilerin amacı diğer sezgisel, metasezgisel optimizasyon tekniklerinde olduğu gibi tüm arama uzayını aramadan, optimal çözüm değerleri bulmaya çalışmaktır.

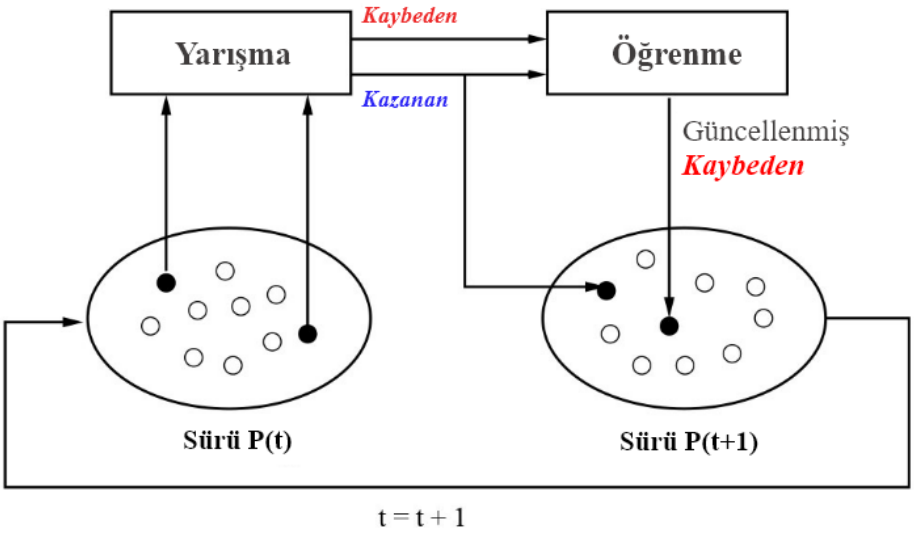
$$v_i = v_x + 2 * rand_{c1} * (pbest_i - x_i) + 2 * rand_{c2} * (gbest_x - x_i) \quad (3.1)$$

$$x_i = x_i + v_x \quad (3.2)$$

Evrimsel bir algoritma olması sebebiyle iterasyonlar halinde ilerleyen bu algoritmada her bir parçacık iterasyon boyunca bulunduğu en iyi konumu ($pbest$) saklar ve bütün sürünün içinde bulunan parçacıklar arasındaki en iyi konumda ($gbest$) ayrıca saklanır. Saklanan bu veriler kullanılarak parçacıkların yani sürünün bir sonraki hareketine karar verilir. Parçacıkların hızlarının güncelleme işlemi Denklem (3.1), konumlarının güncelleme işlemi ise Denklem (3.2)'de gösterilmiştir. Bu işlem istenen yakınsama veya müsaade edilen en fazla iterasyon sayısına ulaşılana kadar devam eder. Bütün bunların amacı optimize edilmesi gereken problemin parametreleri için en iyi yakınsamayı en az sayıda denemeyle sağlamaktır (Kennedy ve Eberhart 1995).

3.2. Rekabetçi Sürü İyileştirici

Cheng ve Jin (2015) tarafından önerilen RSİ algoritması PSO'dan esinlense de konsept olarak farklı olan bir algoritmadır. RSİ'de, sürü $P(t)$ rastgele başlatılan ve yinelemeli olarak güncellenen m adet parçacık içerir. Parçacıklar, küresel veya kişisel en iyi pozisyondan öğrenmekten ziyade rastgele seçilmiş rakiplerden öğrenirler. Her iterasyonda sürüdeki parçacıklar rastgele iki gruba ayrılır $m/2$ ve her gruptaki parçacıklar arasında birebir yarışmalar yapılır. Her yarışmadan sonra kazanan yani en iyi uygunluk değerine sahip olan parçacık doğrudan bir sonraki iterasyona aktarılır. Kaybeden parçacıklar ise kazanan parçacıklardan öğrenerek konum ve hızlarını günceller. Ardından bir sonraki iterasyona aktarılır (Cheng ve Jin 2015). Şekil 3.2'de parçacıkların bir sonraki iterasyona aktarılma işlemi gösterilmiştir.



Şekil 3.2. RSİ'de parçacıkların güncelleme işlemleri (Cheng ve Jin 2015)

Kaybeden parçacığın hız vektörlerinin güncellenmesi işlemi Denklem (3.3)'te gösterilmiştir. $R_1(k,t)$, $R_2(k,t)$, $R_3(k,t)$, [0-1] arasında rastgele seçilmiş vektörlerdir.

$$\begin{aligned} v_{l,k}(t+1) = \quad & R_1(k,t)v_{l,k}(t) \\ & +R_2(k,t)\left(x_{w,k}(t) - x_{l,k}(t)\right) \\ & +\varphi R_3(k,t)\left(\overline{x}_k(t) - x_{l,k}(t)\right) \end{aligned} \tag{3.3}$$

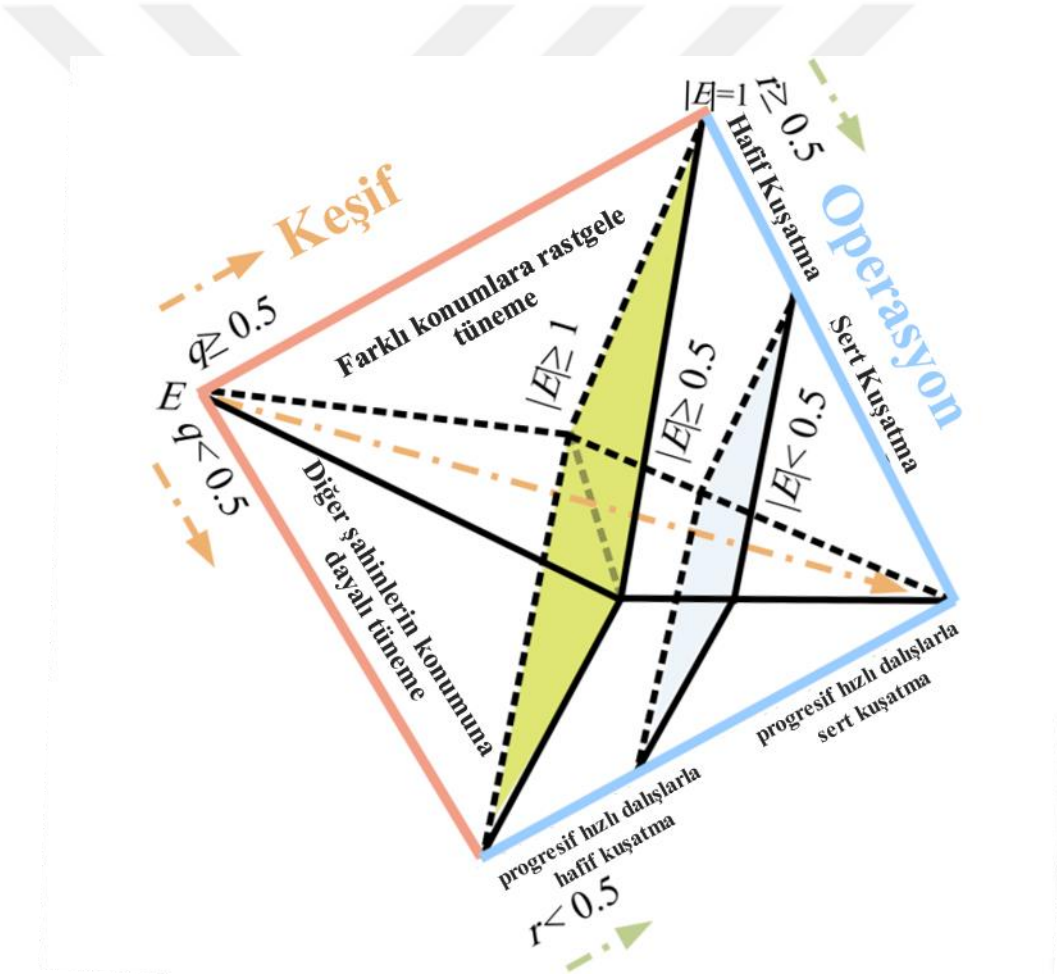
Kaybeden parçacığın konum vektörlerinin güncellenmesi işlemi Denklem (3.4)'te gösterilmiştir.

$$x_{l,k}(t+1) = x_{l,k}(t) + v_{l,k}(t+1) \tag{3.4}$$

Denklem (3.3)'te yer alan $x_{w,k}$ kazanan parçacığın konumunu, $x_{l,k}$ kaybeden parçacığın konumunu, $v_{l,k}$ kaybeden parçacığın hızını belirtmektedir ($k = 1, 2, \ldots, m/2$). $\overline{x}_k(t)$, ilişkili parçacıkların ortalama konum değerini belirtirken, phi (φ) parametresi ise parçacıkların ortalama konum değerinin $\overline{x}(t)$ etki kontrolü için ayarlanması gereken parametredir (Cheng ve Jin 2015, Gu ve ark 2018).

3.3. Harris Şahinleri Optimizasyonu

Heidari ve ark (2019) sundukları sürü zekasına dayalı olan HŞO algoritması bir Harris Şahini'nin diğer şahinlerle iş birliğiyle avlarının peşine düşmesinden ilham alınarak geliştirilmiştir. Henüz literatürde yeni olan bu algoritmanın yapılan çalışmalarla güçlü bir algoritma olduğu gösterilmiş ve farklı uygulamalarda kullanılmıştır (Heidari ve ark 2019). Bu algoritmada şahinler iş birliği içinde farklı yollarla avlarını şaşırtmaya çalışırlar ve saldırırlar. Buna ek olarak, Harris şahinleri avın kaçma modellerine göre kovalamaca türünü seçme yeteneğine sahiptir (Moayedi ve ark 2019). Harris Şahinleri Şekil 3.3.'te verilen keşif ve operasyon aşamalarıyla avlarına tuzak kurar ve harekete geçerler.



Şekil 3.3. HŞO algoritmasında keşif ve operasyon aşamaları grafiği (Heidari ve ark 2019)

3.3.1. Keşif Aşaması

Harris Şahinleri doğaları gereği sahip oldukları güçlü gözleri sayesinde avlarını kolaylıkla tespit ve takip edebilirler, fakat bazen avlarında görülmesi zor olabilir, bu tür durumlarda Harris şahinleri alanı bekler, gözetler ve avını tespit etmek için saatlerce izler. HŞO'da her bir şahin aday çözümdür ve her adımdaki en iyi aday çözüm hedeflenen av veya optimuma yakınsama olarak kabul edilir. Şahinler rastgele konumlarda tüneyip bekleyerek avlarını iki farklı stratejiyle tespit etmeye çalışırlar. Bu stratejilerden (q) ilkinde herhangi bir ava saldırı durumunda hem aile bireylerine hem de ava belirli bir yakınlıkta durabilmektir, bu strateji $q < 0.5$ olarak belirtilir. Diğer strateji ise saldırıya geçecek şahinlerin yuva menzili içinde bulunan rastgele uzun ağaçlara tünemesidir, bu strateji de $q >= 0.5$ olarak belirtilir. Bu iki strateji Denklem (3.5)'te verilmiştir.

$$X(t+1) = \begin{cases} X_{\text{rand}}(t) - r_1 |X_{\text{rand}}(t) - 2r_2 X(t)| & q \geq 0.5 \\ \left(X_{\text{rabbit}}(t) - X_m(t)\right) - r_3\left(LB + r_4(UB - LB)\right) & q < 0.5 \end{cases} \quad (3.5)$$

HŞO algoritması, grubun yuva menzili (LB, UB) içinde rastgele konumlar oluşturmak için iki kuraldan oluşan basit bir model önerir. Birinci kural rastgele konumla diğer şahinlere göre çözüm üretirken, ikinci kural konum farklarından en iyisi ve grubun konum ortalaması değerlerine ek olarak rastgele ölçeklenen değerler içerir. *r3* ölçekleme katsayısı kuralın rassal doğasını artırırken r4 ise 1'e yakın değerler alarak benzer dağılım desenlerinin oluşmasını sağlar. Bu kuralda LB'ye rastgele ölçeklenir bir hareket uzunluğu eklenmektedir. Sonrasında bütün bu bileşiğin çeşitliliğin sağlanması için ve özellik uzayında daha farklı bölgelerin keşfedilmesi için rastgele ölçeklenir başka bir katsayı uygulanır. Aslında daha farklı güncelleme kuralları uygulanabilecekken HŞO algoritmasında şahinlerin davranışlarını taklit edebilecek basit bir kural kullanılmıştır. Denklem (3.6) ile şahinlerin ortalama konumu hesaplanır.

$$X_m(t) = \frac{1}{N} \sum_{i=1}^{N} X_i(t) \quad (3.6)$$

HŞO algoritmasında şahinler yaptıkları keşiften sonra operasyona geçeceklerdir. Operasyona geçtiklerinde ise avın kaçış enerjisine bağlı olarak farklı operasyonel

davranışlarda bulunabilirler. Kaçış esnasında ise avın kaçış enerjisi düşüş gösterir. Bu bağlam Denklem (3.7) ile modellenir.

$$E = 2E_0(1 - \frac{t}{T}) \tag{3.7}$$

Formülde E avın kaçış enerjisini, T maximum iterasyon sayısını, E_0 ise başlangıç enerjisini belirtir. HŞO algoritmasında her iterasyonda E_0 (-1,1) aralığında değişiklik gösterir. Bu değerin 0 dan -1'e doğru değişimi avın (tavşan) fiziksel olarak güçsüzleştiğini gösterirken, 0'dan 1'e doğru değişimi ise avın güçlendiğini gösterir. Dinamik olan kaçış enerjisi E, iterasyonlar arasında düşüş yönelimi gösterir. Kaçış enerjisi $|E| \geq 1$ iken şahinler farklı yerleri keşfetmeye yönelirler, dolayısıyla HŞO keşif fazını işletir. Bunun aksine $|E| < 1$ iken, HŞO bulunan alan (komşuluk) içerisinde operasyon fazına geçip çözüm arayışında bulunacaklar. Kısaca $|E| \geq 1$ iken keşif, $|E| < 1$ iken operasyon fazı işler.

3.3.2. Operasyon Aşaması

Bu fazda Harris şahinleri bir önceki fazda tespit edilerek hedef seçilen ava sürpriz saldırılar düzenlerler. Fakat, avlarda tehlikeli durumlardan kaçmaya çalışacaklardır. Gerçek durumlarda bu şahinlerin farklı takip stratejileri izlemesine sebep olur. HŞO algoritmasının operasyon fazında da bu duruma uyacak şekilde 4 farklı saldırı modeli bulunmaktadır.

Av her zaman tehdit durumlarında kaçmaya çalışacaktır. r değerini avın kaçış başarısı olarak değerlendirilirse, $r < 0.5$ avın başarılı kaçtığı durumu, $r <= 0.5$ te sürpriz saldırı öncesi kaçamayışını temsil etmektedir. Av her ne yaparsa yapsın, şahinler yakalamayı gerçekleştirmek için sert kuşatma veya hafif kuşatma ile cevap verirler. Yani, avın kaçış enerjisi seviyesine göre, avı farklı yönlerden hafifçe veya sertçe kuşatarak çevrelemeye çalışırlar. Gerçek durumlarda şahinler iş birliği içerisinde avlarına sürekli yaklaşarak bir sürpriz saldırır ile sonlandırmak üzere yakalama şansını artırırlar. Sürekli kaçış halinde olan av, bu süreç içerisinde enerji kaybına uğradığı için zamanla yorulur. Avın yorulduğunu gördükçe şahinler ise kuşatmayı yoğunlaştırarak yorgun avı daha kolay yakalama yoluna giderler.

Bu stratejiyi modellemek için HŞO, E parametresini kullanarak hafif ve sert kuşatma arasında geçişler yapar. $|E| \geq 0.5$ iken hafif kuşatma, $|E| < 0.5$ iken ise sert kuşatma uygulanır.

$r \geq 0.5$ ve $|E| \geq 0.5$ iken avın (tavşan) halen kaçmaya çabalayabilecek bir miktar enerjisi vardır. Birkaç yanıltıcı sıçrama yaparak kurtulmaya çalışır fakat sonunda yakalanacaktır. Bu süreçte Harris şahinleri hafif bir şekilde tavşanı çevrelemeye devam ederek, avlarını sürpriz atak öncesi yormaya çabalarlar. Bu davranış Denklem (3.8) ve Denklem (3.9)'da modellendirilmiştir.

$$X(t+1) = \Delta X(t) - E|JX_{\text{rabbit}}(t) - X(t)| \tag{3.8}$$

$$\Delta X(t) = X_{\text{rabbit}}(t) - X(t) \tag{3.9}$$

$\Delta X(t)$ Tavşanın mevcut t iterasyondaki pozisyonu ile pozisyon vektörünün farkını belirtir. $r5$, 0 ila 1 arasında rastgele bir sayıdır, $J = 2(1 - r5)$ ise tavşanın kaçış prosedürü içerisinde rastgele zıplama gücünü temsil eder. J değeri her iterasyonda değişerek tavşanın doğal hareketindeki tahmin edilemezliği taklit eder.

$r \geq 0.5$ ve $|E| < 0.5$ durumlarında av yorulmuştur ve düşük bir kaçış enerjisi kalmıştır. Harris şahinleri ise avlarının bu durumundan faydalanıp sürpriz saldırı ile avlanmayı sonlandırmak için sert bir şekilde avlarını çevrelerler. Bu durumda mevcut pozisyonlar Denklem (3.10) kullanılarak güncellenir.

$$X(t+1) = X_{\text{rabbit}}(t) - E \mid \Delta X(t) \mid \tag{3.10}$$

$|E| \geq 0.5$ olduğu durumda $r < 0.5$ olursa şayet, av (tavşan) halen başarılı bir kaçışı sağlayabilecek enerjiye sahiptir, bu sebeple hafif kuşatma sürpriz atak öncesi hazırlanmıştır. Bu operasyon yöntemi öncekinden daha zekidir.

HŞO içerisinde, bu yöntemde avın kaçışı levy uçuşları (LF) kullanılarak modellenir. LF, avın (özellikle tavşan gibi) zig zag yaparak yanıltıcı ve tahmin edilmesi zor hareketlerle kaçışının ve şahinin hızlı dalışlarla kaçan avın çevresindeki hareketinin taklidi için kullanılır. Hafif kuşatmada, şahinlerin yaptığı sonraki hamle Denklem (3.11) ile modellenmiştir.

$$Y = X_{\text{rabbit}}(t) - E \mid JX_{\text{rabbit}}(t) - X(t) \mid \tag{3.11}$$

Sonrasında ise, olası bu hamle sonucunu bir önceki dalış sonucu ile karşılaştırarak, dalışı iyi veya kötü olarak nitelendirirler. Bu dalışlar Denklem (3.12)'de modellenmiştir.

$$Z = Y + S \times LF(D) \tag{3.12}$$

Burada, D problemin boyutunu, S, 1xD boyutunda rastgele bir vektörü ve LF ise Denklem (3.13) ile hesaplanan levy uçuş fonksiyonunu nitelendirir.

$$LF(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)}} \right)^{\frac{1}{\beta}} \tag{3.13}$$

Formülde u, v 0 ila 1 arasında rastgele değerler, β ise 1,5 olarak belirlenmiş varsayılan bir sabittir. $|E| < 0.5$ ve $r < 0.5$ iken, av(tavşan) yeterli kaçma enerjisine sahip değildir ve sürpriz saldırı öncesi sert kuşatma hazırlanmıştır. Şahinlerin hafif ve sert kuşatmada sırasında pozisyon güncellemeleri Denklem (3.14) ile yapılmaktadır. Sert kuşatma esnasında Y değeri Denklem (3.15)'te Z değeri Denklem (3.12) ile güncellenir.

$$X(t+1) = \begin{cases} Y & \text{if } F(Y) < F\big(X(t)\big) \\ Z & \text{if } F(Z) < F\big(X(t)\big) \end{cases} \tag{3.14}$$

$$Y = X_{\text{rabbit}}(t) - E|JX_{\text{rabbit}}(t) - X_m(t)\,| \tag{3.15}$$

Temel olarak keşif ve operasyon olmak üzere iki aşamadan oluşan popülasyon tabanlı HŞO algoritmasında, Harris Şahinlerinin avları keşfetme, sürpriz saldırı ve farklı saldırı stratejilerinden esinlenilmiştir. HŞO algoritmasının sözde kodu Şekil 3.4'te verilmiştir.

Harris Şahinleri Optimizasyonu Algoritması

Girdiler: Popülasyon boyutu (N) ve maks. yineleme sayısı (T)
Çıktılar: Tavşanın konumu ve uygunluk değeri
Rastgele bir popülasyonla algoritmayı başlat $\boldsymbol{X_i}(\boldsymbol{i = 1, 2, \dots, N})$
while (bitiş kriteri gerçekleşmediği sürece) **do**
  Şahinlerin uygunluk değerini hesapla
  En iyi konumu tavşanın konumu olarak ayarla ($\boldsymbol{X_{tavşan}}$)
  for (her şahin $\boldsymbol{X_i}$ için) **do**
    Tavşanın başlangıç enerjisi $\boldsymbol{E_0}$ ve zıplama gücünü $\boldsymbol{J}$ güncelle
    Kaçış enerjisini ($\boldsymbol{E}$) Denklem (3.7) ile güncelle
    if ($\boldsymbol{|E| \geq 1}$) **then** -- Keşif Aşaması
      Konum vektörünü Denklem (3.5) ile güncelle
    if ($\boldsymbol{|E| < 1}$) **then** -- Operasyon Aşaması
      if ($\boldsymbol{r \geq 0.5\ ve\ |E| \geq 0.5}$) **then** -- *Hafif Kuşatma*
        Konum vektörünü Denklem (3.8) ile güncelle
      else if ($\boldsymbol{r \geq 0.5\ ve\ |E| < 0.5}$) **then** -- *Sert Kuşatma*
        Konum vektörünü Denklem (3.10) ile güncelle
      else if ($\boldsymbol{r < 0.5\ ve\ |E| \geq 0.5}$) **then**-- *Hızlı Dalışla Hafif Kuşatma*
        Konum vektörünü Denklem (3.14) ile güncelle
      else if ($\boldsymbol{r < 0.5\ ve\ |E| < 0.5}$) **then** -- *Hızlı Dalışla Sert Kuşatma*
        Konum vektörünü Denklem (3.14) ile güncelle
Return $\boldsymbol{X_{tavşan}}$

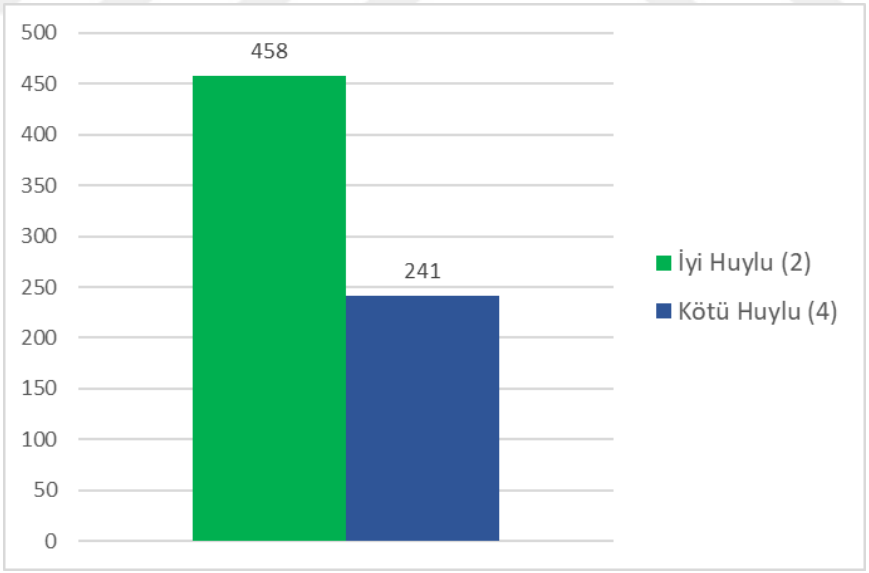
Şekil 3.4. HŞO algoritmasının sözde kodu (Heidari ve ark 2019)

4. ÇALIŞMADA KULLANILAN VERİ SETLERİ

Bu çalışmada geliştirilen modelleri sınıflandırma problemleri üzerinde denemek amacıyla açık erişimli UCI Makine Öğrenmesi Deposundan alınan birbirinden farklı boyutlarda ve girdi sayısında olan popüler veri setleri kullanılmıştır. İkili sınıflandırma problemi olarak göğüs kanseri (Breast Cancer Wisconsin Original) ve diyabet (Pima Indians Diabetes), çoklu sınıflandırma problemi olarak şarap (Wine), iris (Iris) ve tiroit (Thyroid Disease) veri setleri kullanılmıştır.

4.1. Göğüs Kanseri Veri Seti

Göğüs kanseri veri seti 1992 yılında Wisconsin Üniversitesi Hastanesi'nden elde edilmiş verilerden oluşmaktadır. Bu veri seti depodaki diğer göğüs kanseri veri setlerinden farklı olarak (orijinal) şeklinde belirtilmiş versiyonudur (Wolberg 1992). Toplamda 699 örnek içerir. 10 adet öznitelik (birisi örnek kod numarası) ve "İyi Huylu" (2), "Kötü Huylu" (4) olarak iki sınıftan oluşmaktadır. Şekil 4.1.'de veri setinin sınıf bazlı örnek dağılımı verilmiştir.



Şekil 4.1. Göğüs kanseri veri seti sınıf bazlı örnek sayıları

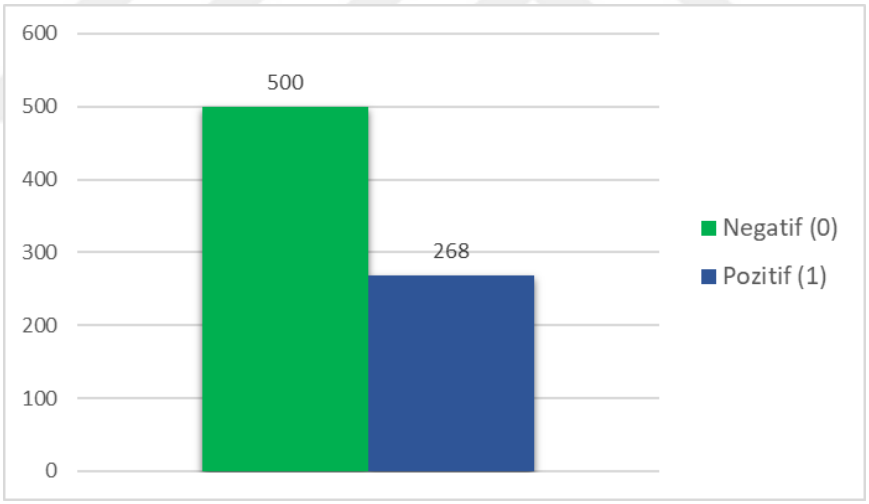
Tablo 4.1.'de göğüs kanseri veri setine ait özniteliklerin değer aralıkları ve ortalama değerleri verilmiştir.

Tablo 4.1. Göğüs kanseri veri setindeki özniteliklerin değer aralıkları ve ortalama değeri

Özellik Adı	Aralık	Ortalama Değer
Kod Numarası (ID)	X	X
Yumru Kalınlığı	[1, 10]	4.4177
Hücre Boyutunun	[1, 10]	3.1345
Hücre Şekli	[1, 10]	3.2074
Marjinal Yapışma	[1, 10]	2.8069
Tek Hücre Boyutu	[1, 10]	3.2160
Çekirdekler	[1, 10]	3.4635
Yumuşak Kromatin	[1, 10]	3.4378
Normal Nükleoid	[1, 10]	2.8670
Mitoz Tipi	[1, 10]	1.5894

4.2. Diyabet Veri Seti

1988 yılında Ulusal Diyabet ve Sindirim ve Böbrek Hastalıkları Enstitüsü'nden alınan diyabet veri setinde amaç 8 farklı özniteliğe göre kişinin diyabet hastası olup olmadığını tespit etmektir (Smith 1988). Pozitif (1) ve negatif (0) olarak iki sınıftan oluşmaktadır. Şekil 4.2.'de veri setinin sınıf bazlı örnek dağılımı verilmiştir.



Şekil 4.2. Diyabet veri seti sınıf bazlı örnek sayıları

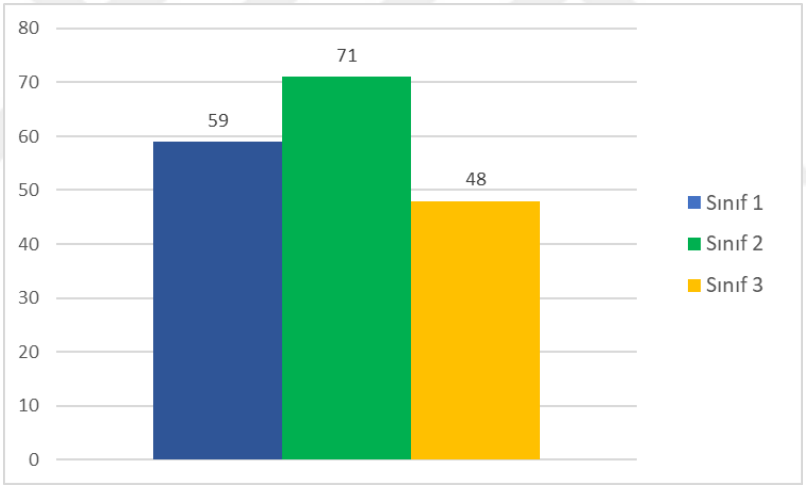
Tablo 4.2.'de diyabet veri setine ait özniteliklerin değer aralıkları ve ortalama değerleri verilmiştir.

Tablo 4.2 Diyabet veri setindeki özniteliklerin değer aralıkları ve ortalama değeri

Özellik Adı	Aralık	Ortalama Değer
Hamilelik Sayısı	[0, 17]	3.845
Glukoz	[0, 199]	120.89
Kan Basıncı (mm Hg)	[0, 122]	69.11
Cilt Kalınlığı (mm)	[0, 99]	20.54
İnsulin (mu U/ml)	[0, 846]	79.80
Vücut kitle indeksi	[0, 67.1]	31.99
Diyabet soyağacı işlevi	[0.08, 2.42]	0.47
Yaş	[21, 81]	33.24

4.3. Şarap Veri Seti

İtalya'da yetiştirilen üç farklı şarapa ait laboratuvar analiz sonuçlarına göre bu veri seti oluşturulmuştur (M. Forina 1991). Bu analizle üç farklı türdeki şarapta bulunan 13 bileşen belirlenmiştir. 178 örnekten oluşan bu veri setinde 13 girdi ve 3 sınıf bulunmaktadır. Şekil 4.3.'te veri setinin sınıf bazlı örnek dağılımı verilmiştir.



Şekil 4.3. Şarap veri seti sınıf bazlı örnek sayıları

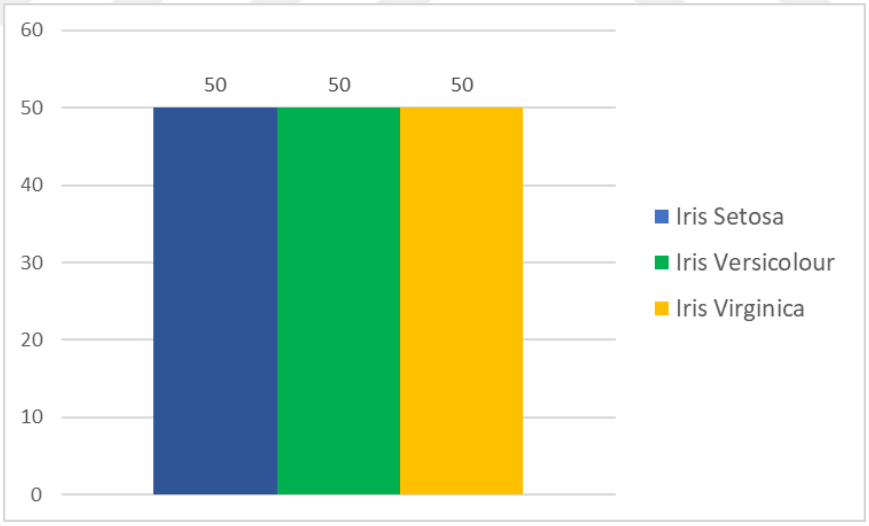
Tablo 4.3.'te şarap veri setine ait özniteliklerin değer aralıkları ve ortalama değerleri verilmiştir.

Tablo 4.3. Şarap veri setindeki özniteliklerin değer aralıkları ve ortalama değeri

Özellik Adı	Aralık	Ortalama Değer
Alkol Oranı	[11.03, 14.83]	13.001
Malik Asit	[0.74, 5.80]	2.3363
Kül	[1.36, 3.23]	2.3365
Kül Alkalisi	[10.60, 30]	19.4949
Magnezyum	[70, 162]	99.7416
Toplam fenoller	[0.98, 3.88]	2.2951
Flavonoid	[0.34, 5.08]	2.0293
Flavonoid olmayan fenoller	[0.13, 0.66]	0.3619
Proantosiyanidin	[0.41, 3.58]	1.5909
Renk Yoğunluğu	[1.28, 13]	5.0581
Renk Tonu	[0.48, 1.71]	0.9574
Seyreltilmiş Şaraptaki OD280/OD315	[1.27, 4]	2.6117
Prolin	[278, 1680]	746.8933

4.4. İris Veri Seti

Literatürde en popüler veri setlerinden olan İris veri seti toplamda 150 veriden oluşmaktadır. İris veri setinde, 4 adet öznitelik ve “Iris Setosa”, “Iris Versicolour” ve “Iris Virginica” olmak üzere 3 adet sınıf bulunmaktadır (Fisher 1988). Şekil 4.4.’te veri setinin sınıf bazlı örnek sayısı verilmiştir.



Şekil 4.4. İris veri seti sınıf bazlı örnek sayıları

Tablo 4.4.’te iris veri setine ait özniteliklerin değer aralıkları ve ortalama değerleri verilmiştir.

Tablo 4.4. İris veri setindeki özniteliklerin değer aralıkları ve ortalama değeri

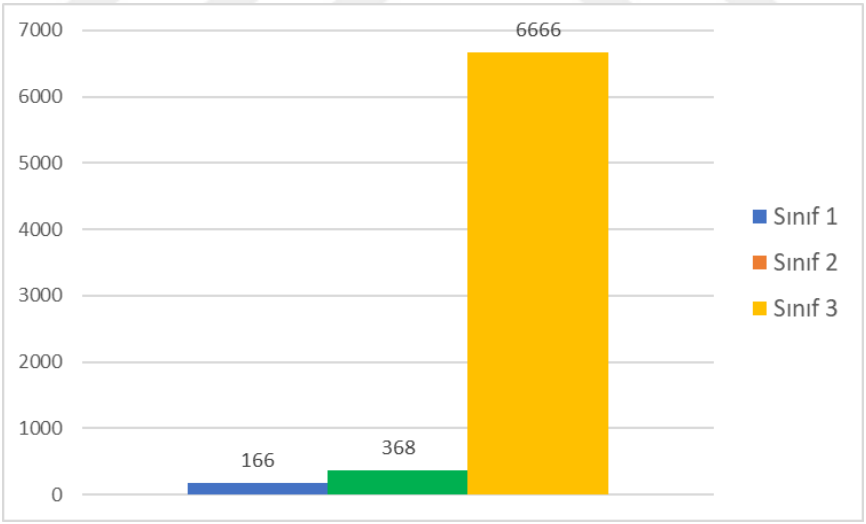
Özellik Adı	Aralık	Ortalama Değer
Çanak Yaprak Uzunluğu (cm)	[4.3, 7.90]	5.8433
Çanak Yaprak Genişliği (cm)	[2.0, 4.40]	3.0540
Taçyaprak Uzunluğu (cm)	[1.0, 6.90]	3.7587
Taçyaprak Genişliği (cm)	[0.1, 2.50]	1.1987

4.5. Tiroit Veri Seti

1987 yılında Garavan Enstitüsünden edinilmiş Tiroit veri seti hastalara ait 15 ikili [0-1], 6 reel sayı olmak üzere 21 girdiden ve teşhis sonucuna göre 3 sınıftan oluşmaktadır:

- 1: Hipertiroid bulunan bireyler
- 2: Hipotiroid bulunan bireyler
- 3: Sağlıklı bireyler (normal kabul edilen)

Veri seti toplam 7200 örnek içermektedir (Quinlan 1987). Şekil 4.5.'te veri setinin sınıf bazlı örnek dağılımı verilmiştir.



Şekil 4.5. Tiroit veri seti sınıf bazlı örnek sayıları

Tablo 4.5'te veri setine ait özniteliklerin değer aralıkları ve ortalama değerleri verilmiştir.

Tablo 4.5. Tiroit veri setindeki özniteliklerin değer aralıkları ve ortalama değeri

Özellik Adı	Aralık	Ortalama Değer
Yaş	[0.01, 0.097]	0.5205
Cinsiyet	[0, 1]	0.3043
Tiroksin	[0, 1]	0.1306
Tiroksin'e ilişkin sorgulama	[0, 1]	0.0154
Anti tiroit ilacı	[0, 1]	0.0128
Hasta	[0, 1]	0.0383
Hamile	[0, 1]	0.0108
Tiroit Cerrahisi	[0, 1]	0.014
I131 tedavi	[0, 1]	0.0168
Hipotiroidizm	[0, 1]	0.0656
Hipertiroidizm	[0, 1]	0.0688
Lityum	[0, 1]	0.0126
Guatr	[0, 1]	0.0082
Tümör	[0, 1]	0.0256
Hipopitüiter	[0, 1]	0.0001
Psikolojik	[0, 1]	0.0489
TSH	[0.0, 0.53]	0.0049
T3	[0.00005, 0.18]	0.02
TT4	[0.002, 0.6]	0.1094
T4U	[0.017, 0.233]	0.0978
FTI	[0.002, 0.642]	0.1132

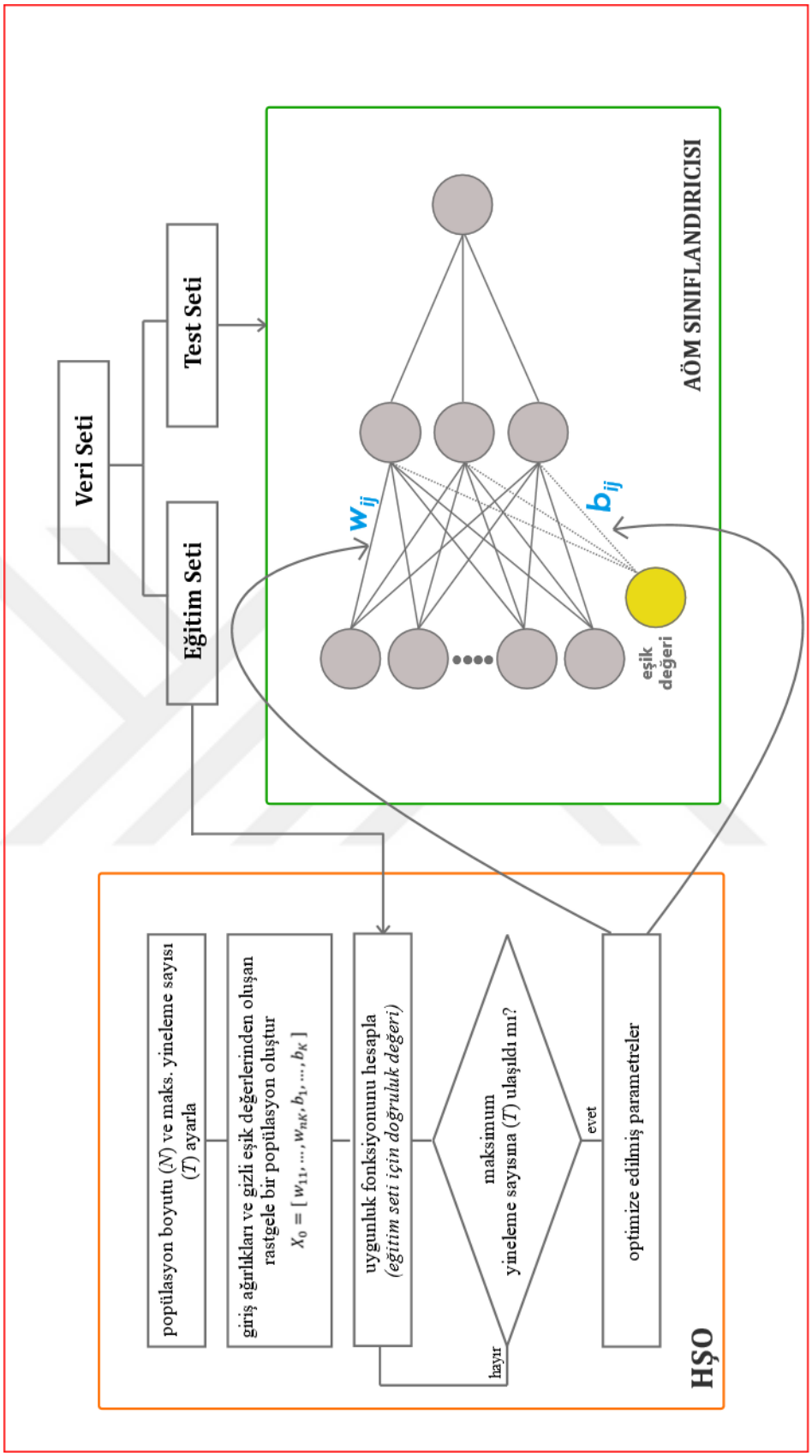
5. SINIFLANDIRMA İÇİN ÖNERİLEN HŞO-AÖM YÖNTEMİ

AÖM'de giriş ağırlıkları ve gizli eşik değerlerinin rastgele seçilmesi, bu parametrelerin ayarlanması için gereken zamandan tasarruf sağlamaktadır. Ancak çıktı ağırlıkları önceden ayarlanmış giriş ağırlıkları ve gizli eşik değerine göre hesaplandığı için gereksiz giriş ağırlıkları ve gizli eşik değerleri sınıflandırıcının kötü sonuç vermesine neden olabilir.

Bu çalışmada sunulan HŞO-AÖM modeliyle genel anlamda AÖM'nin giriş ağırlıkları ve gizli eşik değerlerinin en uygununun seçilerek TGKİB'nin genelleştirme performansının artırılması sağlanmıştır. Bu modelin çalışmasına ait detaylar 5 adımda özetlenebilir:

1. Parametre olarak alınan popülasyon sayısı ve problem boyutuna göre gizli eşik değeri ve giriş ağırlığından oluşan Harris Şahinlerinin konumunun bulunduğu [-1,1] (Han ve ark 2013, Niu ve ark 2016, Eshtay ve ark 2018) aralığında rastgele bir popülasyon oluşturulur.
2. Giriş ağırlıkları ve gizli eşik değerlerinden oluşan her bir şahine karşılık gelen çıktı ağırlıkları, yineleme olmaksızın Denklem (2.8)'de yer alan Moore-Penrose ile hesaplanır. En iyi konum tavşanın konumu olarak ayarlanır. Her şahine ait konumun uygunluk değeri Denklem (5.1)'e göre değerlendirilir.
3. Şahinlerin konumu keşif aşamasında iki strateji ile güncellenir. Denklem (3.5)'te verilmiş olan bu iki strateji diğer aile üyelerine göre tüneme ile güncelleme ($q < 0.5$) ve saldırıya geçecek şahinlerin yuva menzili (tavşanın konumu) içinde bulunan rastgele uzun ağaçlara tünemeleriyle güncelleme ($q \geq 0.5$) işlemidir.
4. Keşif aşamasından sonra gelen operasyon aşamasında 4 farklı strateji ile şahinlerin konumu güncellenir. Bunlar, Denklem (3.8)'de verilen hafif kuşatma, Denklem (3.10)'da verilen sert kuşatma ve Denklem (3.14)'te verilen hızlı dalışlarla hafif ve sert kuşatma yöntemleridir.
5. Birinci adım hariç diğer adımlar belirtilen maksimum iterasyon sayısı kadar tekrar edilir ve optimizasyon işlemlerinden elde edilen en iyi çıktı ağırlıkları modelin genelleştirme performansını ölçmek için test veri setine uygulanır.

Modelin öğrenme parametrelerinin optimize edilmesi işleminin grafiksel gösterimi Şekil 5.1'de verilmiştir.



Şekil 5.1. HŞO-AÖM modeliyle parametrelerin optimizasyonu

5.1. Performans Değerlendirme Kriterleri

Bu çalışmada önerilen modelin değerlendirilmesi için kullanılan sınıflandırma performans ölçütlerinden doğruluk Denklem (5.1)'de, hassasiyet Denklem (5.2)'de, seçicilik Denklem (5.3)'te, G-mean Denklem (5.4)'te, F-score Denklem (5.5)'te, standart sapma Denklem (5.6)'da verilmiştir.

$$Doğruluk = \frac{DP + DN}{DP + DN + HP + HN} \; x\,100 \tag{5.1}$$

Doğruluk ölçütü sınıflandırma sistemlerinde en sık kullanılan değerlendirme kriteridir. Denklem (5.1), Denklem (5.2) ve Denklem (5.3)'te yer alan DP, pozitif olan sınıfta doğru kategorize edilmiş, DN ise negatif sınıfta doğru kategorize edilmiş değerlerdir. HP, pozitif olan sınıfta hatalı kategorize edilmiş, HN ise negatif olan sınıfta hatalı olarak kategorize edilmiş değerlerdir.

$$Hassasiyet = \frac{DP}{DP + HN} \tag{5.2}$$

Hassasiyet ölçütü aynı zamanda DPO (doğru pozitif oranı) olarak ta bilinmektedir. Bir modelin normal olmayan kayıtları ne kadar iyi tanıyabildiğini gösterir. Yüksek hassasiyet iyi sınıflandırıcı anlamına gelmektedir.

$$Seçicilik = \frac{DN}{DN + TP} \tag{5.3}$$

Seçicilik ölçütü aynı zamanda DNO (doğru negatif oranı) olarak ta bilinmektedir. Bir sınıflandırma da önerilmiş yöntemin normal durumları tanımlama yeteneğini gösterir. Seçicilik değeri, duyarlılık gibi yüksek olursa ve farkları çok azsa ($\% < 1$), sınıflandırıcı çok daha iyidir (Mohapatra ve ark 2015).

$$G - mean = \sqrt{Hassasiyet * Seçicilik} \tag{5.4}$$

G-mean iki sınıfın performansının tek bir metrikte birleştirilmesidir. Hem hassasiyeti hem de seçiciliği dengeli bir şekilde ölçer. Eğer iki sınıftan birinin doğruluğu çok düşükse G-mean değeri yüksek olamaz (Mohapatra ve ark 2015).

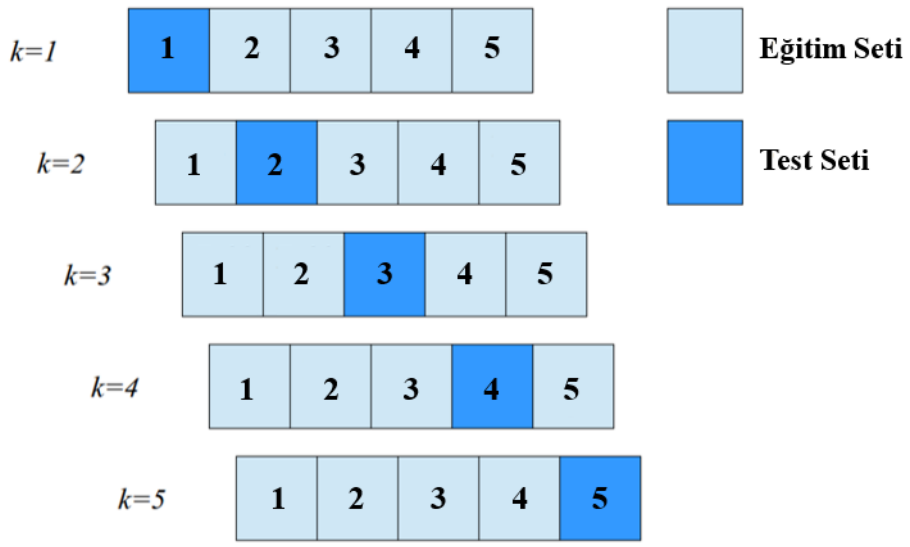
$$F - score = (2 * Hassasiyet * Seçicilik) / (Hassasiyet + Seçicilik) \quad (5.5)$$

F-score'da tıpkı G-mean gibi iki sınıfın performansını ölçmek için kullanılan önemli metriklerdendir. Denklem (5.6)'da formülü verilen standart sapma (Std) N farklı değerin ortalama değere göre dağılımını ifade eder.

$$Std = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \overline{x})^2}{N}} \quad (5.6)$$

Bir diğer performans ölçütü ise alıcı işletim karakteristiği (ROC) eğrisidir. ROC eğrisinde x ekseni yanlış pozitif oranını, y ekseni doğru pozitif oranını belirtir. Area under the curve (AUC), ROC eğrisinin altında kalan alanı ifade etmektedir. AUC değerinin yüksek olması sınıflandırıcının performansının iyi olduğu gösterir (Raghuwanshi ve Shukla 2018).

Çapraz doğrulama, bir modelin bir veri setine ait genelleme yeteneğini değerlendirmek için kullanılan tekniklerdendir. Şekil 5.2.'de 5 katlamalı çapraz doğrulama yöntemine ait gösterim verilmiştir.



Şekil 5.2. 5-katlamalı çapraz doğrulama yöntemi (Saporetti ve ark 2018)

Çapraz doğrulama teknikleri arasından *k-katlamalı* çapraz doğrulama en ünlü olanıdır. (Saporetti ve ark 2018).

5.2. Uygunluk Fonksiyonu

Metasezgisel algoritmalarda üretilen her çözümün uygunluğunu ölçebilmek için uygunluk fonksiyonu kullanılması gerekmektedir. Bu çalışmada evrimsel sinir ağlarında genellikle uygulanmış olan (de Oliveira ve Ludermir 2012, Zhang ve ark 2015, Eshtay ve ark 2018), Denklem (5.7)'de görüldüğü gibi yanlış sınıflandırma oranının minimizasyonu seçilmiştir.

$$Uygunluk = min\left(1 - \frac{1}{N}\sum_{i=1}^{N} I(f(x_i) = y_i)\right) \quad (5.7)$$

Denklem (5.7)'de bulunan $f(x_i)$, eğitim örneklerinin (N) tahmin edilen değeri, y_i ise bu örneklerin gerçek değerleridir. $I(doğru) = 1$, doğru sınıflandırmayı, $I(yanlış) = 0$ ise yanlış sınıflandırmayı belirtir.

5.3. Parametrelerin Belirlenmesi

Optimizasyon modellerinin gizli katmandaki az sayıda nöron sayısıyla daha iyi sonuç verme durumunun araştırılması amacıyla nöron sayısı 5'den 50'ye kadar her seferinde kademeli olarak 5'er artırılmıştır (Pacifico ve Ludermir 2013, Mohapatra ve ark 2015). Optimize edilmeyen AÖM için yüksek nöron sayılarında başarım performansını gözlemlemek amacıyla 50'den 300'e kadar 50'şer artırılmıştır. AÖM'de gizli katman çıkış matrisini hesaplamak için sigmoid, sin, triangular, hard limit ve radial basis fonksiyonları aktivasyon fonksiyonu olarak kullanılmaktadır. Ancak bu çalışmada kullanılacak dört model (AÖM, RSİ-AÖM, PSO-AÖM ve önerilen HŞO-AÖM) için doğrusal olmayışından dolayı, sinir ağlarında sıklıkla kullanılan sigmoid fonksiyon [$f(x) = 1/1 + \exp(-x)$] seçilmiştir.

Önerilen ve karşılaştırılacak optimizasyon modellerinin eğitiminde maksimum iterasyon sayısı 100 (Sánchez-Monedero ve ark 2010, Matias ve ark 2013, Eshtay ve ark 2018), popülasyon sayısı 50 olarak alınmıştır. Optimizasyon algoritmalarına ait diğer parametreler Tablo 5.1.'de verilmiştir.

Tablo 5.1. Optimizasyon algoritmalarının parametreleri

Algoritma	Parametre	Değer
PSO	c1	2
	c2	2
	Maks. iterasyon	100
RSİ	phi	0.03, 0.05
	Maks. iterasyon	100
HŞO	Maks. iterasyon	100

Çalışmada kullanılan veri setlerinde tüm girişler (özellikler) [0,1] aralığında min-maks normalize işleminden geçirilmiştir. Bu işleme ait formül Denklem 5.8'de verilmiştir.

$$y_i = \frac{x_i - min(x)}{max(x) - min(x)} \tag{5.8}$$

Denklemde y_i normalize edilmiş veriyi, x_i ise verinin normalize edilmeden önceki değerini ifade etmektedir.

Gerçekleştirilen her denemede modellerin eğitimi ve testi için kullanılan veriler literatürde sıklıkla kullanılan (Saraswathi ve ark 2011, Saporetti ve ark 2018) 5-fold çapraz doğrulama kullanılarak ayrılmıştır. Tablo 5.2'de veri setlerinin özellikleri ve toplam örnek sayıları verilmiştir.

Tablo 5.2. Çalışmada kullanılan veri setlerine ait özellikler

Veri Seti	Öznitelikler	Sınıf Sayısı	Örnek Sayısı
Diyabet	8	2	768
Göğüs Kanseri	10	2	699
İris	4	3	150
Şarap	13	3	178
Tiroit	21	3	7200

Deneysel çalışmanın geçerliliğini sağlamak için her veri setinde toplamda 50 çalıştırmanın ortalaması alınarak sonuçlar kaydedilmiştir.

6. DENEYSEL SONUÇLARIN İNCELENMESİ

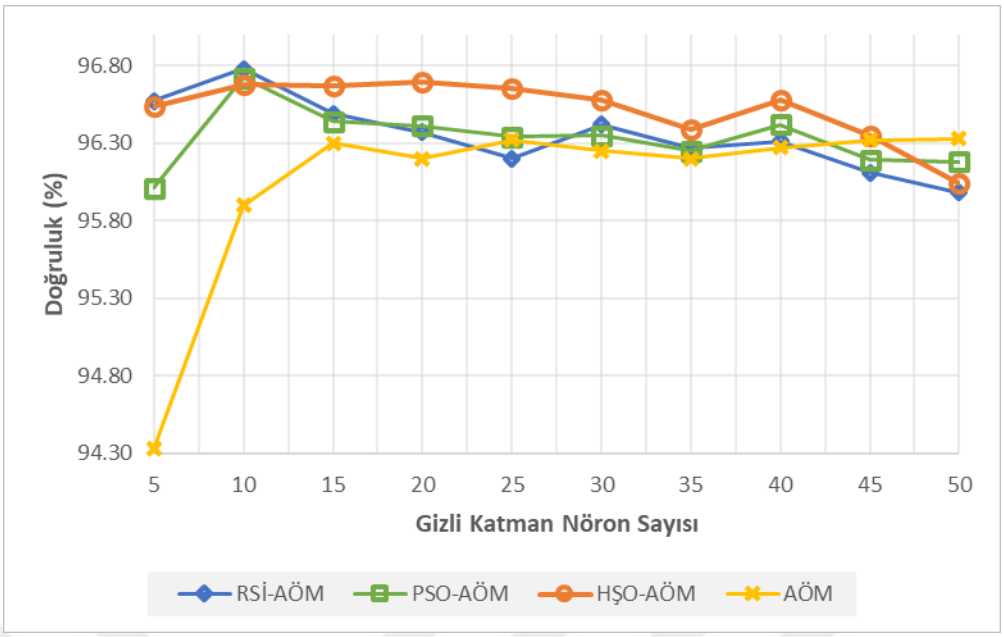
Bu bölümde önerilen HŞO-AÖM modelinin performansı, literatürde popülasyon tabanlı en popüler meta sezgisel optimizasyon algoritmalarından olan PSO algoritmasıyla oluşturulmuş PSO-AÖM modeli, geçtiğimiz yıllarda sunulmuş olan sürü zekasına dayalı RSİ algoritması ile oluşturulmuş RSİ-AÖM modeli ve temel AÖM modeli ile karşılaştırmalı olarak ölçülmüştür. Yapılan deneylere ait sonuçlar ikili ve çoklu sınıflandırma problemleri başlıkları altında verilmiş ardından modellerin literatürdeki diğer çalışmalarla karşılaştırması yapılmıştır. Çalışmaya ait tüm denemeler i7 6700HQ işlemcili kişisel bilgisayarda, MATLAB ortamında yapılmıştır.

6.1. İkili Sınıflandırma Problemlerinde Önerilen Model Başarımı

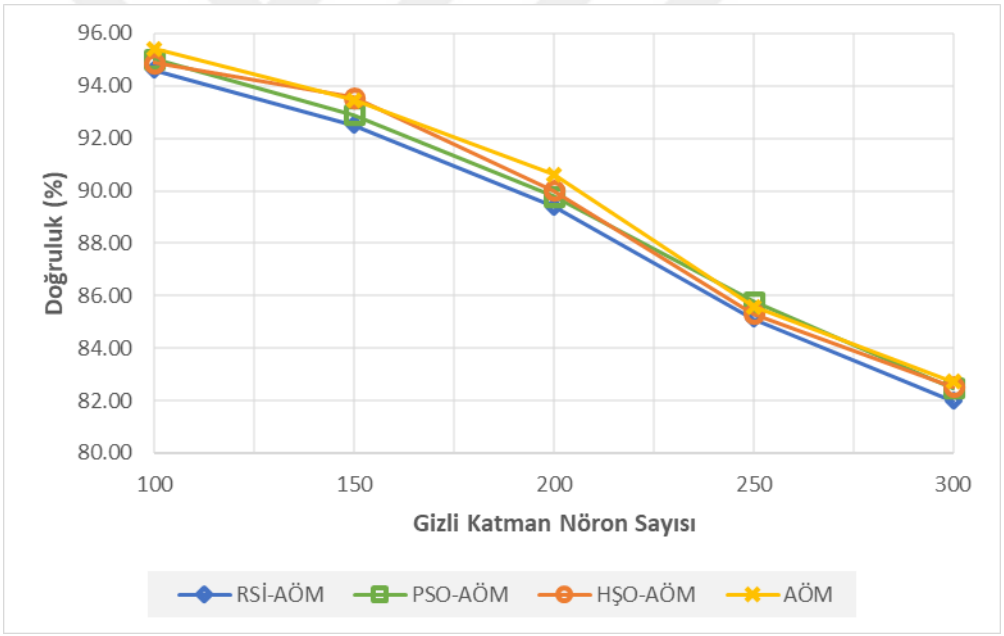
Modellerin ikili sınıflandırma problemlerindeki performansı sınıfların veri setindeki dağılımı dengeli olmayan göğüs kanseri ve diyabet veri setleri kullanılarak ölçülmüştür.

6.1.1. Göğüs Kanseri Veri Seti Sonuçları

Şekil 6.1 ve Şekil 6.2'de önerilen HŞO-AÖM modeli ile PSO-AÖM, RSİ-AÖM ve temel AÖM modellerinin test aşamasındaki ortalama yüzde doğruluk değerlerinin gizli katmandaki artan nöron sayısıyla değişim grafikleri verilmiştir. Şekil 6.1'de görüldüğü üzere gizli katmandaki nöron sayısı 5 olduğunda optimizasyon modellerinin ortalama yüzde doğruluk değerleri birbirlerine yakınken temel AÖM modeli görece daha kötü bir sonuç vermiştir. HŞO-AÖM modelinin diğer modellere göre az farkla üstün olduğu yine bu grafikten görülmektedir. Şekil 6.2'de tüm modellerin ortalama yüzde doğruluk değerlerinin gizli katmandaki 100 nörondan sonra düşüşe geçtiği yani modellerin artan nöron sayısıyla daha kötü sonuç vermeye başladığı görülmektedir.



Şekil 6.1. Modellerin göğüs kanseri veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi



Şekil 6.2. Modellerin göğüs kanseri veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi

Tablo 6.1.'de göğüs kanseri veri seti için gizli katmandaki nöron sayısına bağlı olarak modellerin toplam 50 çalıştırmada almış olduğu maksimum yüzde doğruluk değerleri, ortalama yüzde doğruluk değerleri ve doğruluk değerlerinin standart sapması (Std.) verilmiştir. Optimizasyon modellerinden önerilen HŞO-AÖM modeli gizli katmanda bulunan 15 nöronla 0.00388 standart sapma ile maksimum %97.42 doğruluk değerine, RSİ-AÖM modeli gizli katmanda bulunan 5 nöronla 0.00437 standart sapma

ile maksimum %97.28 doğruluk değerine ve PSO-AÖM modeli gizli katmanda bulunan 5 nöronla 0.00599 standart sapma ile maksimum %97.28 doğruluk değerine ulaşırken, temel AÖM modeli ise gizli katmanda bulunan 15 nöronla 0.00466 standart sapma ile maksimum %96.85 doğruluk değerine ulaşmıştır.

Tablo 6.1. Modellerin göğüs kanseri veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri

M	HŞO-AÖM			RSİ-AÖM			PSO-AÖM			AÖM		
	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std
5	96.85	96.54	.00241	**97.28**	96.57	.00437	**97.28**	96.31	.00599	95.85	94.33	.00911
10	96.99	**96.68**	.00148	97.14	**96.78**	.00246	97.28	**96.72**	.00319	96.42	95.91	.00271
15	**97.42**	96.67	.00388	97.00	96.49	.00311	96.85	96.44	.00297	**96.85**	96.31	.00466
20	97.14	96.68	.00344	96.85	96.37	.00296	97.14	96.41	.00333	96.85	96.21	.00433
25	96.86	96.65	.00217	97.00	96.20	.00378	96.71	96.34	.00254	96.85	**96.32**	.00278
30	97.28	96.58	.00290	96.85	96.42	.00350	96.85	96.35	.00311	96.71	96.25	.00293
35	97.00	96.39	.00421	97.14	96.27	.00359	96.71	96.25	.00316	96.71	96.21	.00352
40	96.86	96.58	.00238	96.71	96.31	.00285	96.85	96.42	.00370	96.71	96.28	.00302
45	96.71	96.35	.00272	97.00	96.21	.00444	96.85	96.19	.00428	96.71	96.32	.00235
50	96.71	96.04	.00270	96.42	95.98	.00353	96.42	96.18	.00243	96.57	96.32	.00214
100	95.42	94.88	.00381	95.28	94.61	.00432	95.71	95.02	.00442	96.00	95.42	.00345
150	94.57	93.56	.00711	93.42	92.52	.00726	94.14	92.90	.00683	94.56	93.46	.00650
200	91.56	90.01	.00832	91.28	89.41	.00796	90.98	89.80	.00983	92.13	90.62	.00891
250	88.27	85.33	.01256	86.69	85.12	.00975	87.55	85.79	.01168	87.27	85.59	.01285
300	83.98	82.53	.01009	83.27	81.99	.00984	84.27	82.48	.01370	85.69	82.72	.01277

Tablo 6.2.'de göğüs kanseri veri seti için modellerin toplamda 50 çalışma sonucunda en iyi ortalama yüzde doğruluk değeri veren gizli nörona ait sonuçlar verilmiştir. Standart sapma Std. olarak belirtilmiştir. Tabloyu incelediğimizde göğüs kanseri veri setinde optimizasyon modellerinin başarımının birbirine yakın olduğu görülmektedir. RSİ-AÖM modeli 0.00246 standart sapma ve %96.78 ortalama doğruluk değeri ile az farkla diğer modellerden daha iyi bir başarım elde etmiştir. PSO-AÖM modeli 0.00599 standart sapma ve %96.72 ortalama doğruluk değeri ile sıralamada ikinci, HŞO-AÖM modeli ise 0.00148 standart sapma ve %96.69 ortalama doğruluk değeri ile üçüncü gelmiştir. Karşılaştırması yapılan dört model arasından en düşük başarıma sahip model 0.00278 standart sapma ve %96.32 ortalama doğruluk değeri ile temel AÖM modeli olmuştur. Maksimum doğruluk oranı olarak baktığımızda %97.42 değeriyle HŞO-AÖM modeli üstün gelmiştir. Minimum test doğruluk oranını ise %95.99 değeri ile temel AÖM modeli göstermiştir.

Tablo 6.2. Göğüs kanseri veri seti için modellerin çalışmasından alınan sonuçlar

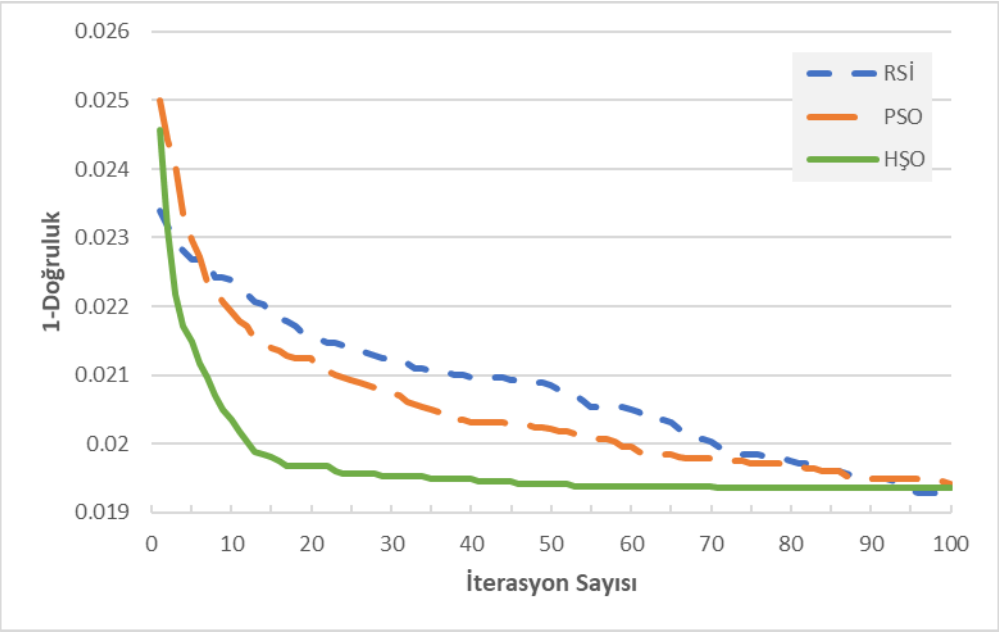
Birim		**HŞO-AÖM**	**RSİ-AÖM**	**PSO-AÖM**	**AÖM**
Test Doğruluk Oranı	Ortalama (%)	96.68	**96.78**	96.72	96.32
	En İyi (%)	97.42	97.28	97.28	96.85
	En Kötü (%)	96.13	96.28	96.28	95.99
	Std	.00148	.00246	.00599	.00278
Gizli Katman Nöron Sayısı		**10**	**10**	**10**	50
Eğitim Süresi (sn)		17.7134	5.6082	7.9643	**0.0128**

Gizli katmanda bulunan nöron sayısı olarak baktığımızda HŞO-AÖM, RSİ-AÖM ve PSO-AÖM modellerinin gizli katmanında bulunan 10 nöron ile temel AÖM modeline göre daha az nöron sayısında en iyi ortalama doğruluk değerini aldıkları görülmektedir. Tablo 6.3.'te göğüs kanseri veri seti için gizli katmandaki nöron sayısına bağlı olarak modellerin toplam 50 çalıştırmada almış olduğu hassasiyet, seçicilik, g-mean ve f-score değerleri verilmiştir. Tüm modellerin birbirine yakın sonuçlar aldığı görülmektedir.

Tablo 6.3. Göğüs kanseri veri seti için modellerin hassasiyet, seçicilik, g-mean ve f-score sonuçları

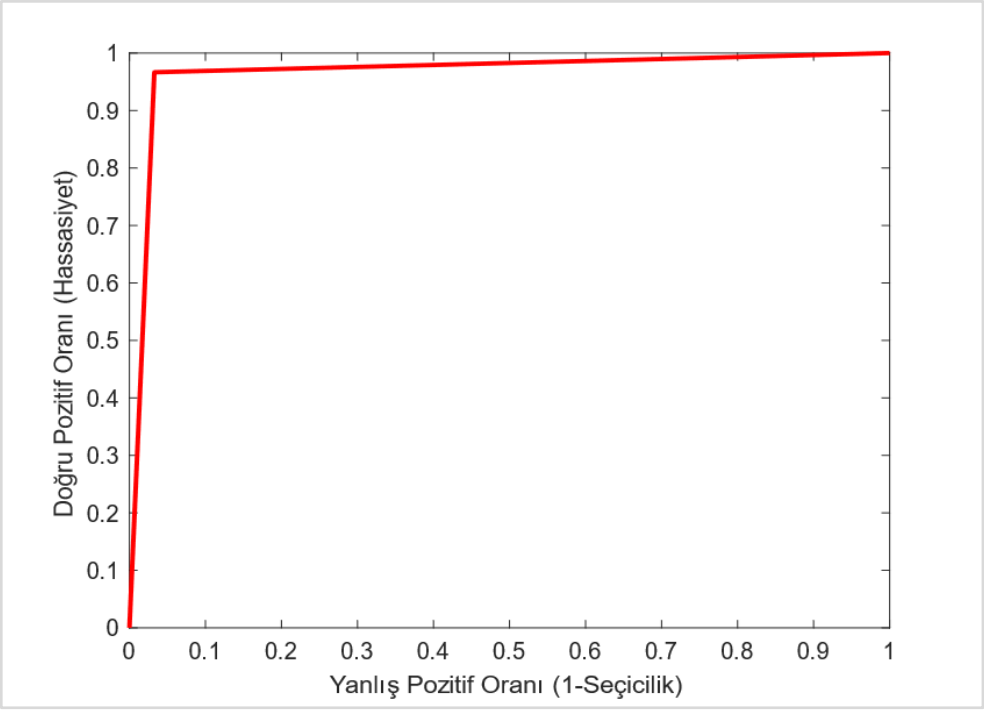
M	**HŞO-AÖM**				**RSİ-AÖM**				**PSO-AÖM**				**AÖM**			
	Hassasiyet	Seçicilik	G mean	F score	Hassasiyet	Seçicilik	G mean	F score	Hassasiyet	Seçicilik	G mean	F score	Hassasiyet	Seçicilik	G mean	F score
5	.9635	.9688	.9660	.9733	.9629	.9709	.9667	.9735	.9646	.9602	.9621	.9716	.9742	.8846	.9276	.9577
10	.9653	.9697	.9674	.9744	.9648	.9734	.9690	.9751	.9659	.9697	.9676	.9747	.9729	.9328	.9524	.9689
15	.9640	.9719	.9678	.9743	.9666	.9619	.9640	.9730	.9673	.9589	.9629	.9726	.9708	.9485	.9595	.9718
20	.9666	.9672	.9668	.9744	.9672	.9568	.9618	.9721	.9684	.9560	.9620	.9724	.9703	.9465	.9581	.9710
25	.9662	.9672	.9666	.9742	.9688	.9490	.9586	.9709	.9673	.9560	.9615	.9719	.9692	.9519	.9603	.9719
30	.9675	.9626	.9648	.9737	.9697	.9540	.9616	.9726	.9690	.9531	.9608	.9721	.9699	.9485	.9590	.9713
35	.9659	.9601	.9628	.9723	.9690	.9506	.9596	.9714	.9695	.9494	.9592	.9713	.9695	.9482	.9586	.9710
40	.9677	.9622	.9648	.9737	.9699	.9502	.9598	.9717	.9710	.9515	.9609	.9726	.9701	.9489	.9593	.9716
45	.9668	.9572	.9618	.9720	.9699	.9473	.9583	.9710	.9716	.9435	.9573	.9710	.9718	.9469	.9591	.9719
50	.9668	.9481	.9572	.9697	.9701	.9402	.9549	.9694	.9709	.9444	.9574	.9708	.9712	.9482	.9594	.9719
100	.9594	.9287	.9436	.9608	.9710	.8989	.9338	.9594	.9701	.9124	.9404	.9624	.9729	.9187	.9452	.9654
150	.9577	.8936	.9246	.9514	.9659	.8477	.9043	.9443	.9675	.8561	.9095	.9471	.9697	.8680	.9170	.9512
200	.9642	.7785	.8657	.9269	.9633	.7627	.8565	.9228	.9638	.7729	.8622	.9255	.9640	.7963	.8756	.9309
250	.9587	.6531	.7895	.8956	.9592	.6462	.7858	.8944	.9577	.6686	.7988	.8985	.9618	.6547	.7920	.8976
300	.9566	.5760	.7408	.8778	.9555	.5622	.7313	.8744	.9568	.5739	.7389	.8777	.9585	.5775	.7423	.8792

Şekil 6.3.'te göğüs kanseri veri seti için çalışmada kullanılan optimizasyon algoritmalarının eğitim sürecindeki yakınsama grafikleri verilmiştir.



Şekil 6.3. Göğüs kanseri veri seti için modellerin yakınsama grafikleri

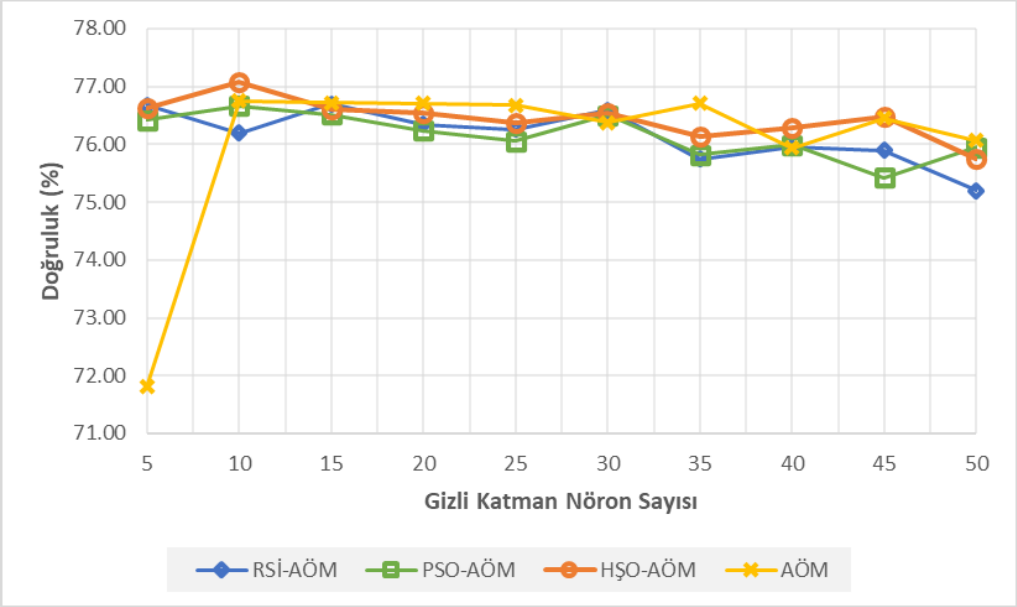
Şekil 6.4.'te önerilen HŞO-AÖM modelinin ROC eğrisi verilmiştir. HŞO-AÖM, RSİ-AÖM, PSO-AÖM ve AÖM modellerinin AUC değerleri sırasıyla 0.9679, 0.9691, 0.9678 ve 0.9605 olarak belirlenmiştir.



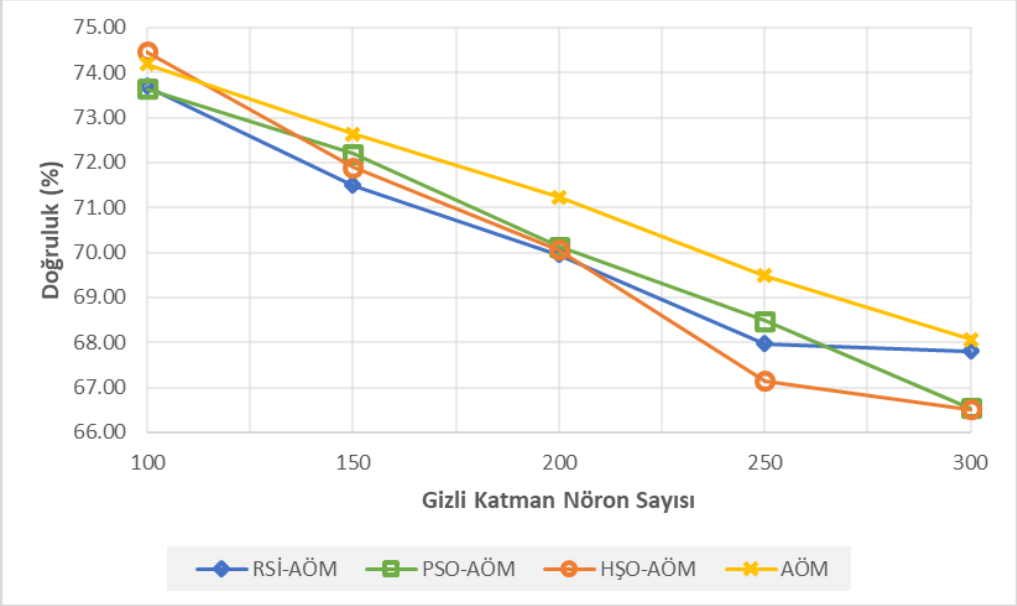
Şekil 6.4. HŞO-AÖM modelinin göğüs kanseri veri seti üzerinde ROC eğrisi

6.1.2. Diyabet Veri Seti Sonuçları

Şekil 6.5 ve Şekil 6.6'da önerilen HŞO-AÖM modeli ile PSO-AÖM, RSİ-AÖM ve temel AÖM modellerinin test aşamasındaki ortalama yüzde doğruluk değerlerinin gizli katmandaki artan nöron sayısıyla değişim grafikleri verilmiştir. Şekil 6.5'de görüldüğü üzere gizli katmandaki nöron sayısı 5 olduğunda optimizasyon modellerinin ortalama yüzde doğruluk değerleri birbirlerine yakınken, temel AÖM modeli görece daha kötü bir sonuç vermiştir.



Şekil 6.5. Modellerin diyabet veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi



Şekil 6.6. Modellerin diyabet veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi

Şekil 6.6'da tüm modellerin ortalama yüzde doğruluk değerlerinin gizli katmandaki 100 nörondan sonra düşüşe geçtiği yani modellerin artan nöron sayısıyla daha kötü sonuç vermeye başladığı görülmektedir. Tablo 6.4.'te diyabet veri seti için gizli katmandaki nöron sayısına bağlı olarak modellerin toplam 50 çalıştırmada almış olduğu maksimum yüzde doğruluk değerleri, ortalama yüzde doğruluk değerleri ve doğruluk değerlerinin standart sapması (Std.) verilmiştir. Optimizasyon modelleri arasında RSİ-AÖM modeli gizli katmanda bulunan 35 nöronla 0.00909 standart sapma ile maksimum %78.51 doğruluk değerine, önerilen HŞO-AÖM modeli gizli katmanda bulunan 10 nöronla 0.00818 standart sapma ile maksimum %78.39 doğruluk değerine, ve PSO-AÖM modeli gizli katmanda bulunan 30 nöronla 0.00909 standart sapma ile maksimum %77.99 doğruluk değerine ulaşırken, temel AÖM modeli ise gizli katmanda bulunan 10 nöronla 0.00679 standart sapma ile maksimum %77.99 doğruluk değerine ulaşmıştır.

Tablo 6.4. Modellerin diyabet veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri

M	HŞO-AÖM			RSİ-AÖM			PSO-AÖM			AÖM		
	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std
5	77.60	76.62	.00549	77.47	76.68	.00579	77.21	76.42	.00607	74.08	71.81	.01466
10	**78.39**	**77.07**	.00818	77.21	76.20	.00607	77.74	**76.67**	.00616	**77.99**	**76.75**	.00679
15	78.12	76.60	.00948	77.47	**76.70**	.00509	77.61	76.52	.00749	77.87	76.72	.00687
20	77.47	76.55	.00587	77.48	76.34	.00736	76.95	76.24	.00589	77.48	76.71	.00424
25	77.86	76.38	.00876	77.21	76.26	.00573	77.73	76.06	.01100	77.87	76.68	.00818
30	77.60	76.54	.00602	**78.51**	76.59	.00909	**77.99**	76.50	.00909	77.09	76.38	.00459
35	77.47	76.14	.00832	76.43	75.74	.00573	76.83	75.82	.00797	77.87	76.71	.00895
40	76.82	76.28	.00602	77.74	75.95	.00973	76.82	75.99	.00517	77.22	75.94	.00974
45	77.22	76.47	.00496	77.22	75.90	.01145	76.56	75.42	.00537	77.60	76.45	.00619
50	76.96	75.74	.00613	75.91	75.20	.00623	77.47	75.95	.00612	77.86	76.07	.00707
100	75.79	74.47	.00784	74.74	73.70	.00825	75.38	73.64	.00830	75.66	74.19	.00895
150	73.30	71.90	.00723	73.69	71.50	.01075	73.18	72.20	.00469	73.96	72.64	.00797
200	72.40	70.07	.01642	71.49	69.95	.01359	72.40	70.13	.01353	72.79	71.24	.00839
250	70.57	68.76	.01369	70.31	68.67	.00901	69.26	67.89	.00785	70.84	69.48	.01056
300	67.83	66.51	.01089	69.15	67.81	.01040	67.59	66.54	.00584	69.79	68.07	.01077

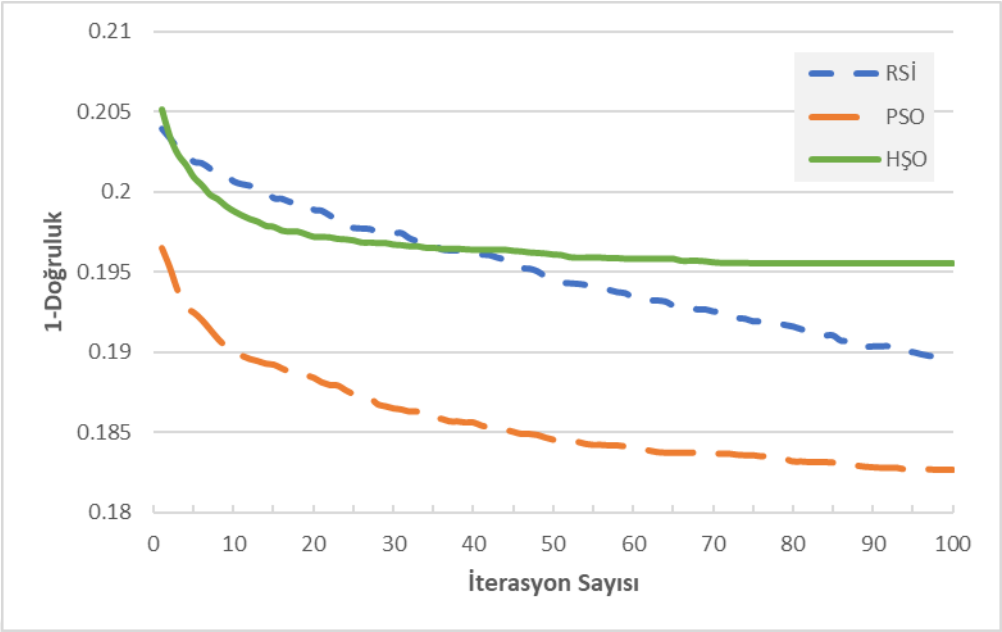
Tablo 6.5.'te diyabet veri seti için modellerin toplamda 50 çalışma sonucunda en iyi ortalama yüzde doğruluk değeri veren gizli nörona ait sonuçlar verilmiştir. Standart sapma Std. olarak belirtilmiştir. Tabloyu incelediğimizde diyabet veri setinde tüm modellerin başarımının birbirine yakın olduğu görülmektedir. HŞO-AÖM modeli 0.00818 standart sapma ve %77.07 ortalama doğruluk değeri ile az farkla diğer

modellerden daha iyi bir başarım elde etmiştir. Temel AÖM modeli 0.00679 standart sapma ve %76.75 ortalama doğruluk değeri ile sıralamada ikinci, RSİ-AÖM modeli ise 0.00509 standart sapma ve %76.70 ortalama doğruluk değeri ile üçüncü gelmiştir. Karşılaştırması yapılan dört model arasından en düşük başarıma sahip model 0.00616 standart sapma ve %76.67 ortalama doğruluk değeri ile PSO-AÖM modeli olmuştur. Maksimum doğruluk oranı olarak baktığımızda %78.51 değeriyle RSİ-AÖM modeli üstün gelmiştir. Minimum test doğruluk oranını ise %75.78 değeri ile PSO-AÖM ve temel AÖM modelleri vermiştir.

Tablo 6.5. Diyabet veri seti için modellerin çalışmasından alınan sonuçlar

Birim		**HŞO-AÖM**	**RSİ-AÖM**	**PSO-AÖM**	**AÖM**
Test Doğruluk Oranı	Ortalama (%)	**77.07**	76.70	76.67	76.75
	En İyi (%)	78.39	78.51	77.99	77.99
	En Kötü (%)	75.79	75.91	75.78	75.78
	Std	.00818	.00509	.00616	.00679
Gizli Katman Nöron Sayısı		10	15	10	10
Eğitim Süresi (sn)		22.1422	6.2993	12.3265	**0.0075**

Gizli katmanda bulunan nöron sayısı olarak baktığımızda HŞO-AÖM, PSO-AÖM ve temel AÖM modellerinin gizli katmanında bulunan 10 nöron ile RSİ-AÖM modeline göre daha az nöron sayısında en iyi ortalama doğruluk değerlerini aldıkları görülmektedir. Şekil 6.7.'de diyabet veri seti için çalışmada kullanılan optimizasyon algoritmalarının eğitim sürecindeki yakınsama grafikleri verilmiştir.



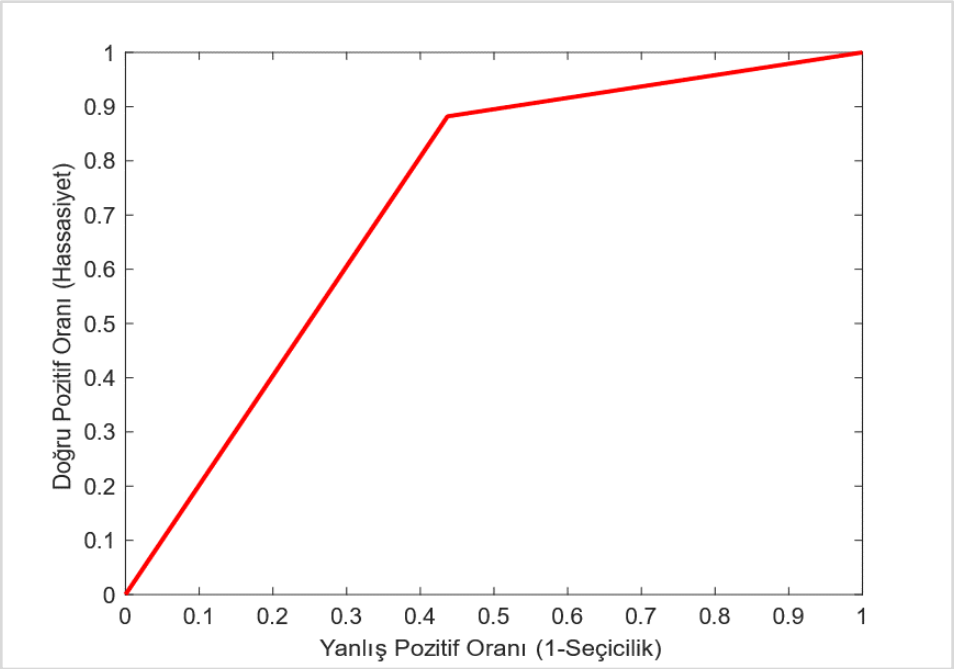
Şekil 6.7. Diyabet veri seti için modellerin yakınsama grafikleri

Tablo 6.6.'da göğüs kanseri veri seti için gizli katmandaki nöron sayısına bağlı olarak modellerin toplam 50 çalıştırmada almış olduğu hassasiyet, seçicilik, g-mean ve f-score değerleri verilmiştir.

Tablo 6.6. Diyabet veri seti için modellerin hassasiyet, seçicilik, g-mean ve f-score sonuçları

M	HŞO-AÖM				RSİ-AÖM				PSO-AÖM				AÖM			
	Hassasiyet	Seçicilik	G mean	F score	Hassasiyet	Seçicilik	G mean	F score	Hassasiyet	Seçicilik	G mean	F score	Hassasiyet	Seçicilik	G mean	F score
5	.8794	.5548	.6972	.8304	.8840	.5481	.6945	.8314	.8826	.5435	.6911	.8298	.8876	.4017	.5889	.8042
10	.8820	.5630	.7032	.8336	.8768	.5478	.6922	.8273	.8790	.5571	.6986	.8306	.8844	.5497	.6961	.8319
15	.8804	.5524	.6958	.8304	.8802	.5559	.6982	.8311	.8826	.5462	.6929	.8302	.8820	.5531	.6972	.8313
20	.8804	.5511	.6948	.8301	.8820	.5423	.6898	.8291	.8782	.5463	.6914	.8277	.8852	.5467	.6945	.8318
25	.8784	.5501	.6938	.8289	.8826	.5388	.6882	.8288	.8786	.5402	.6872	.8268	.8856	.5452	.6937	.8317
30	.8756	.5598	.6991	.8292	.8812	.5507	.6952	.8304	.8828	.5452	.6923	.8302	.8814	.5445	.6906	.8293
35	.8782	.5436	.6896	.8272	.8740	.5398	.6857	.8243	.8742	.5419	.6870	.8246	.8802	.5560	.6982	.8310
40	.8772	.5492	.6925	.8279	.8774	.5395	.6866	.8260	.8814	.5332	.6843	.8269	.8742	.5452	.6888	.8255
45	.8810	.5478	.6930	.8297	.8722	.5479	.6896	.8248	.8688	.5402	.6839	.8214	.8802	.5486	.6937	.8295
50	.8740	.5401	.6858	.8242	.8656	.5400	.6824	.8195	.8744	.5451	.6891	.8256	.8768	.5441	.6891	.8266
100	.8458	.5559	.6843	.8114	.8456	.5344	.6703	.8071	.8386	.5458	.6748	.8052	.8490	.5421	.6769	.8107
150	.8236	.5240	.6553	.7923	.8188	.5212	.6515	.7889	.8270	.5262	.6581	.7947	.8284	.5363	.6645	.7975
200	.8002	.5151	.6400	.7764	.8040	.5045	.6352	.7767	.8108	.4970	.6334	.7790	.8154	.5203	.6498	.7865
250	.7814	.5125	.6310	.7647	.7846	.5042	.6272	.7650	.7844	.4822	.6131	.7603	.7918	.5137	.6363	.7711
300	.7704	.4686	.5993	.7493	.7692	.5080	.6233	.7566	.7586	.4914	.6087	.7467	.7738	.5070	.6245	.7588

Şekil 6.8.'de önerilen HŞO-AÖM modelinin ROC eğrisi verilmiştir. HŞO-AÖM, RSİ-AÖM, PSO-AÖM ve AÖM modellerinin AUC değerleri sırasıyla 0.7225, 0.7180, 0.7180 ve 0.7181 olarak belirlenmiştir.



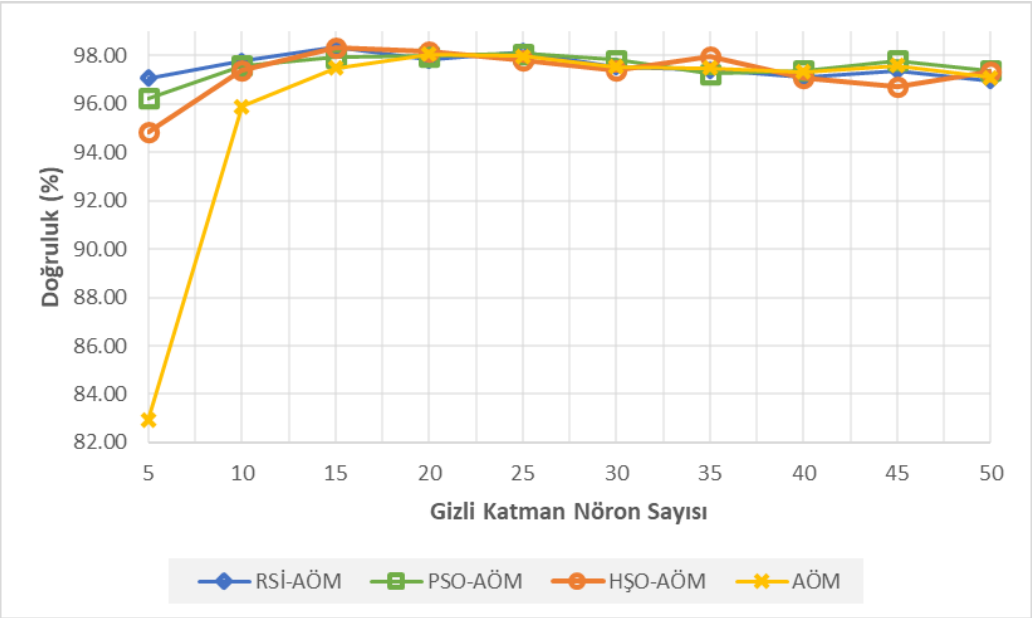
Şekil 6.8. HŞO-AÖM modelinin diyabet veri seti üzerinde ROC eğrisi

6.2. Çoklu Sınıflandırma Problemlerinde Önerilen Model Başarımı

Modellerin çoklu sınıflandırma problemlerindeki performansı Şarap, İris ve Tiroit veri setleri kullanılarak ölçülmüştür.

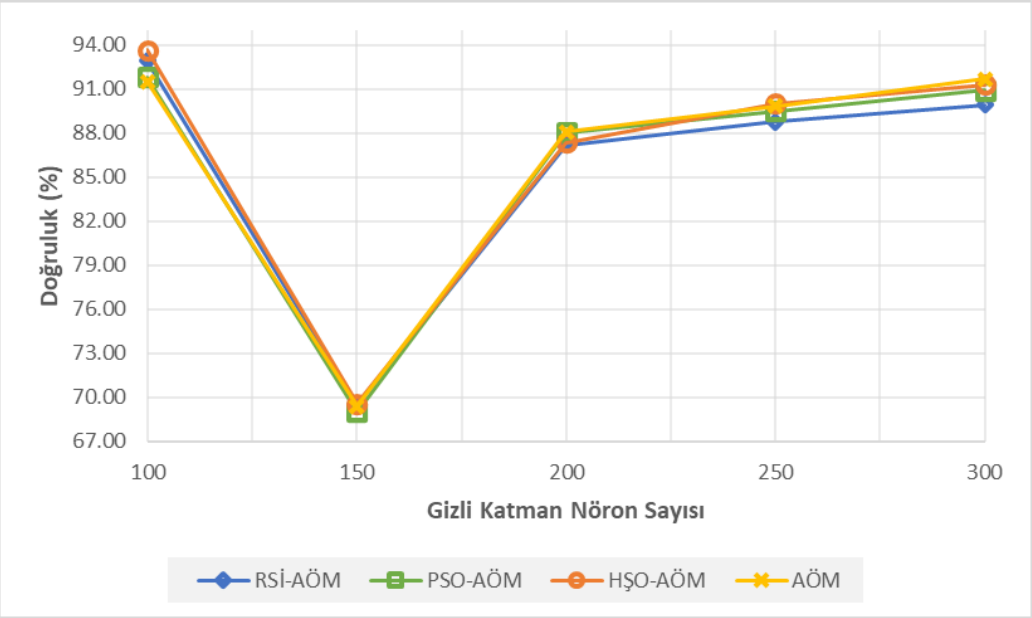
6.2.1. Şarap Veri Seti için Sonuçlar

Şekil 6.9 ve Şekil 6.10'da önerilen HŞO-AÖM modeli ile PSO-AÖM, RSİ-AÖM ve temel AÖM modellerinin test aşamasındaki ortalama yüzde doğruluk değerlerinin gizli katmandaki artan nöron sayısıyla değişim grafikleri verilmiştir. Şekil 6.9'da görüldüğü üzere gizli katmandaki nöron sayısı 5 olduğunda optimizasyon modellerinin ortalama yüzde doğruluk değerleri birbirlerine yakınken, temel AÖM modeli görece daha kötü bir sonuç vermiştir.



Şekil 6.9. Modellerin şarap veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi

Şekil 6.10'da tüm modellerin ortalama yüzde doğruluk değerlerinin gizli katmandaki 100 nörondan sonra düşüşe geçtiği, 150 nörondan sonra kırılıma uğrayıp tekrardan yükseldiği görülmektedir. Tablo 6.7.'de şarap veri seti için gizli katmandaki nöron sayısına bağlı olarak modellerin toplam 50 çalıştırmada almış olduğu maksimum yüzde doğruluk değerleri, ortalama yüzde doğruluk değerleri ve doğruluk değerlerinin standart sapması (Std.) verilmiştir. Tüm modeller maksimum %100 doğruluk değerine ulaşmıştır.



Şekil 6.10. Modellerin şarap veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi

Tablo 6.8.'de şarap veri seti için modellerin toplamda 50 çalışma sonucunda en iyi ortalama yüzde doğruluk değeri veren gizli nörona ait sonuçlar verilmiştir. Standart sapma Std. olarak belirtilmiştir. Tabloyu incelediğimizde şarap veri setinde tüm modellerin başarımının birbirine yakın olduğu görülmektedir.

Tablo 6.7. Modellerin şarap veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri

M	HŞO-AÖM			RSİ-AÖM			PSO-AÖM			AÖM		
	Maks. (%)	Ort. (%)	Std.	Maks. (%)	Ort. (%)	Std.	Maks. (%)	Ort. (%)	Std.	Maks. (%)	Ort. (%)	Std.
5	96.06	94.83	.01084	98.87	97.07	.00994	98.87	97.58	.00845	88.76	82.91	.03813
10	98.86	97.41	.00839	99.44	97.76	.00754	94.90	91.79	.01570	97.78	95.89	.01940
15	98.89	**98.31**	.00750	**100.00**	**98.32**	.01473	98.87	97.91	.00661	98.33	97.47	.01145
20	**100.00**	98.14	.01029	98.87	97.81	.00721	76.46	69.03	.05495	**100.00**	**98.04**	.00923
25	98.32	97.81	.00665	99.44	98.15	.00879	**100.00**	97.97	.01218	99.44	97.98	.00999
30	100.00	97.41	.01002	98.33	97.53	.00883	90.49	88.04	.01849	98.87	97.52	.00801
35	100.00	97.98	.01068	99.44	97.41	.01258	98.89	**98.08**	.00765	99.43	97.47	.01438
40	98.89	97.08	.01452	99.44	97.08	.01473	92.67	89.49	.02023	98.33	97.31	.00792
45	98.32	96.69	.01394	100.00	97.36	.01610	100.00	97.82	.01072	98.89	97.59	.00705
50	98.33	97.31	.00832	98.89	96.96	.01197	93.84	90.90	.01740	98.33	97.09	.00825
100	95.52	93.65	.01865	95.52	92.92	.01690	98.87	97.25	.00933	93.30	91.51	.01634
150	74.16	69.60	.03435	75.87	69.49	.04575	98.89	97.36	.00930	72.98	69.42	.02542
200	91.57	87.37	.02156	91.56	87.20	.02417	98.32	97.80	.00551	91.03	88.10	.01874
250	92.16	90.00	.01687	92.71	88.77	.02386	97.21	96.23	.00831	91.02	89.82	.01434
300	94.37	91.28	.02029	92.11	89.93	.01832	98.89	97.37	.00795	94.95	91.69	.01427

RSİ-AÖM modeli 0.01473 standart sapma ve %98.32 ortalama doğruluk değeri ile az farkla diğer modellerden daha iyi bir başarım elde etmiştir. HŞO-AÖM modeli 0.00750 standart sapma ve %98.31 ortalama doğruluk değeri ile sıralamada ikinci, PSO-AÖM modeli ise 0.00765 standart sapma ve %98.08 ortalama doğruluk değeri ile üçüncü gelmiştir. Karşılaştırması yapılan dört model arasından en düşük başarıma sahip model 0.00923 standart sapma ve %98.04 ortalama doğruluk değeri ile AÖM modeli olmuştur. Maksimum doğruluk oranı olarak baktığımızda %100 değeriyle RSİ-AÖM ve AÖM modelleri üstün gelmiştir. Minimum test doğruluk oranını ise %94.95 değeri ile RSİ-AÖM modeli vermiştir.

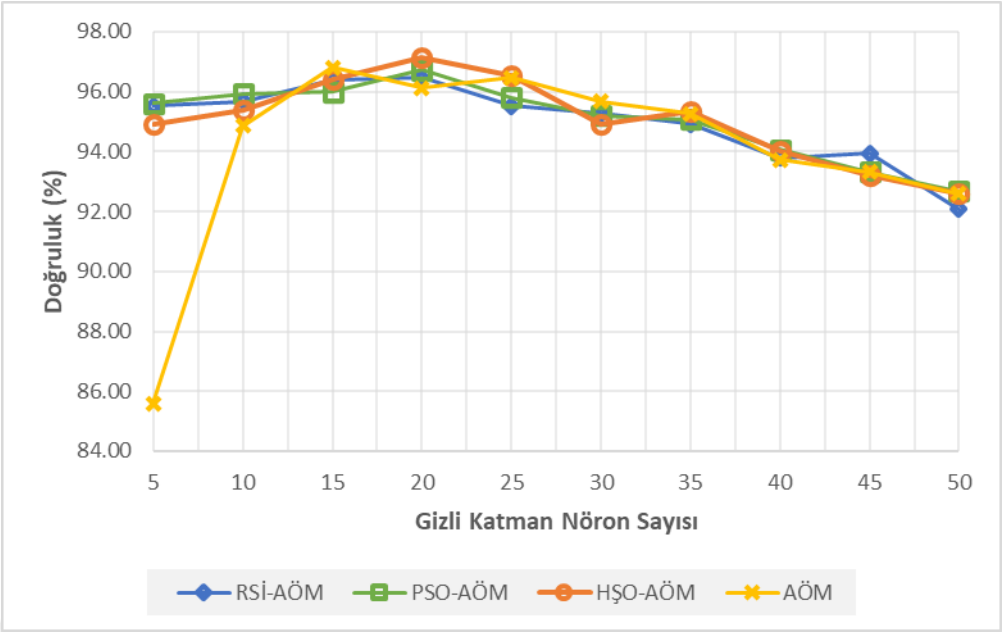
Tablo 6.8. Şarap veri seti için modellerin çalışmasından alınan sonuçlar

Birim		HŞO-AÖM	RSİ-AÖM	PSO-AÖM	AÖM
Test Doğruluk Oranı	Ortalama (%)	98.31	**98.32**	98.08	98.04
	En İyi (%)	98.89	100.00	98.89	100.00
	En Kötü (%)	97.19	94.95	96.62	96.65
	Std	.00750	.01473	.00765	.00923
Gizli Katman Nöron Sayısı		15	15	25	20
Eğitim Süresi (sn)		8.5595	2.3184	6.1216	**0.0025**

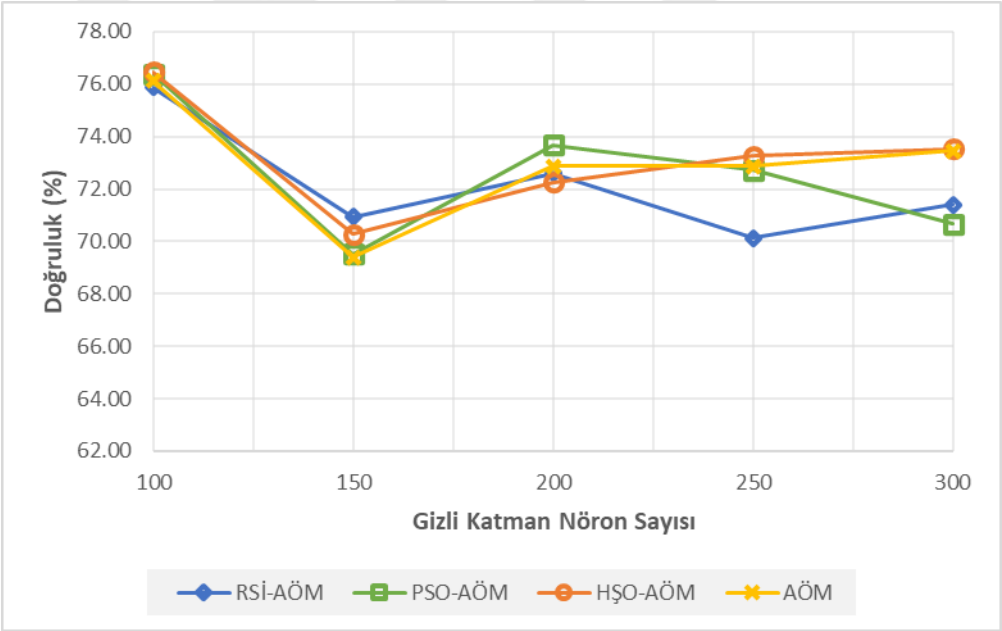
Gizli katmanda bulunan nöron sayısı olarak baktığımızda RSİ-AÖM ve HŞO-AÖM modellerinin gizli katmanda bulunan 15 nöron sayısı ile PSO-AÖM ve AÖM modellerine göre daha az nöron sayısında en iyi ortalama doğruluk değerlerini aldıkları görülmektedir.

6.2.2. İris Veri Seti için Sonuçlar

Şekil 6.11. ve Şekil 6.12.'de önerilen HŞO-AÖM modeli ile PSO-AÖM, RSİ-AÖM ve temel AÖM modellerinin test aşamasındaki ortalama yüzde doğruluk değerlerinin gizli katmandaki artan nöron sayısıyla değişim grafikleri verilmiştir. Şekil 6.11.'de görüldüğü üzere gizli katmandaki nöron sayısı 5 olduğunda optimizasyon modellerinin ortalama yüzde doğruluk değerleri birbirlerine yakınken, temel AÖM modeli görece daha kötü bir sonuç vermiştir. Şekil 6.12'de tüm modellerin ortalama yüzde doğruluk değerlerinin gizli katmandaki 100 nörondan sonra düşüşe geçtiği, 150 nörondan sonra kırılıma uğrayıp tekrardan yükseldiği görülmektedir.



Şekil 6.11. Modellerin iris veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi



Şekil 6.12. Modellerin iris veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi

Tablo 6.9.'da iris veri seti için gizli katmandaki nöron sayısına bağlı olarak modellerin toplam 50 çalıştırmada almış olduğu maksimum yüzde doğruluk değerleri, ortalama yüzde doğruluk değerleri ve doğruluk değerlerinin standart sapması (Std.) verilmiştir. Optimizasyon modelleri arasında önerilen HŞO-AÖM modeli gizli katmanda bulunan 5 nöronla 0.02089 standart sapma ile maksimum %98.67 doğruluk değerine, RSİ-AÖM modeli gizli katmanda bulunan 20 nöronla 0.01442 standart sapma ile maksimum %98 doğruluk değerine, ve PSO-AÖM modeli gizli katmanda bulunan

10 nöronla 0.01489 standart sapma ile maksimum %98 doğruluk değerine ulaşırken, temel AÖM modeli ise gizli katmanda bulunan 15 nöronla 0.01326 standart sapma ile maksimum %98.67 doğruluk değerine ulaşmıştır.

Tablo 6.9. Modellerin iris veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri

M	HŞO-AÖM			RSİ-AÖM			PSO-AÖM			AÖM		
	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std
5	**98.67**	94.93	.02089	96.67	95.53	.00632	96.67	95.60	.00783	88.67	85.60	.02136
10	96.67	95.40	.01016	97.33	95.67	.00903	**98.00**	95.93	.01489	96.67	94.87	.01407
15	98.00	96.40	.01142	97.33	96.40	.01052	98.00	96.00	.01176	**98.67**	**96.80**	.01326
20	98.67	**97.13**	.00996	**98.00**	**96.47**	.01442	98.00	**96.73**	.00798	98.00	96.13	.00820
25	98.67	96.53	.01398	97.33	95.53	.01259	98.00	95.80	.01259	98.00	96.47	.01219
30	96.00	94.93	.02136	96.67	95.27	.00858	98.00	95.20	.01326	98.00	95.67	.01186
35	96.67	95.33	.01176	97.33	94.93	.01838	96.67	95.07	.01052	96.67	95.27	.01063
40	95.33	94.00	.00770	96.00	93.80	.01807	96.00	94.07	.01554	96.67	93.73	.01412
45	94.67	93.20	.01326	95.33	93.93	.01016	94.67	93.33	.01474	95.33	93.33	.01217
50	94.00	92.60	.01016	93.33	92.07	.01489	94.00	92.67	.01257	95.33	92.60	.01489
100	80.67	76.47	.02668	82.00	75.87	.03553	82.00	76.40	.03488	82.00	76.13	.03910
150	74.00	70.27	.03695	74.00	70.93	.02633	73.33	69.53	.02267	74.67	69.40	.03855
200	77.33	72.27	.04039	76.67	72.60	.03671	77.33	73.67	.03067	77.33	72.87	.03048
250	78.00	73.27	.02232	76.00	70.13	.04155	78.67	72.73	.04222	78.00	72.87	.03297
300	80.00	73.53	.03611	78.00	71.40	.03070	75.33	70.67	.03158	78.67	73.47	.03186

Tablo 6.10.'da iris veri seti için modellerin toplamda 50 çalışma sonucunda en iyi ortalama yüzde doğruluk değeri veren gizli nörona ait sonuçlar verilmiştir. Standart sapma Std. olarak belirtilmiştir. Tabloyu incelediğimizde iris veri setinde tüm modellerin başarımının birbirine yakın olduğu görülmektedir. HŞO-AÖM modeli 0.00996 standart sapma ve %97.13 ortalama doğruluk değeri ile az farkla diğer modellerden daha iyi bir başarım elde etmiştir. Temel AÖM modeli 0.01326 standart sapma ve %96.8 ortalama doğruluk değeri ile sıralamada ikinci, PSO-AÖM modeli ise 0.00798 standart sapma ve %96.73 ortalama doğruluk değeri ile üçüncü gelmiştir. Karşılaştırması yapılan dört model arasından en düşük başarıma sahip model 0.01442 standart sapma ve %96.47 ortalama doğruluk değeri ile RSİ-AÖM modeli olmuştur. Maksimum doğruluk oranı olarak baktığımızda %98.67 değeriyle HŞO-AÖM ve AÖM modelleri üstün gelmiştir. Minimum test doğruluk oranını ise %94.00 değeri ile RSİ-AÖM ve AÖM modelleri vermiştir.

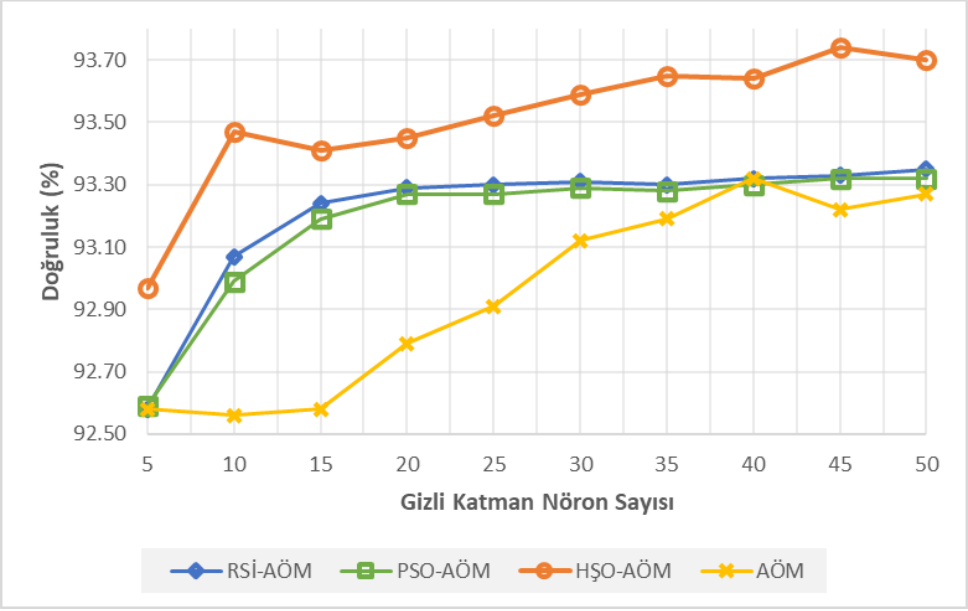
Tablo 6.10. İris veri seti için modellerin çalışmasından alınan sonuçlar

Birim		HŞO-AÖM	RSİ-AÖM	PSO-AÖM	AÖM
Test Doğruluk Oranı	Ortalama (%)	**97.13**	96.47	96.73	96.8
	En İyi (%)	98.67	98.00	98.00	98.67
	En Kötü (%)	96.00	94.00	96.00	94.00
	Std	.00996	.01442	.00798	.01326
Gizli Katman Nöron Sayısı		20	20	20	**15**
Eğitim Süresi (sn)		9.7308	2.5849	5.4405	**0.0034**

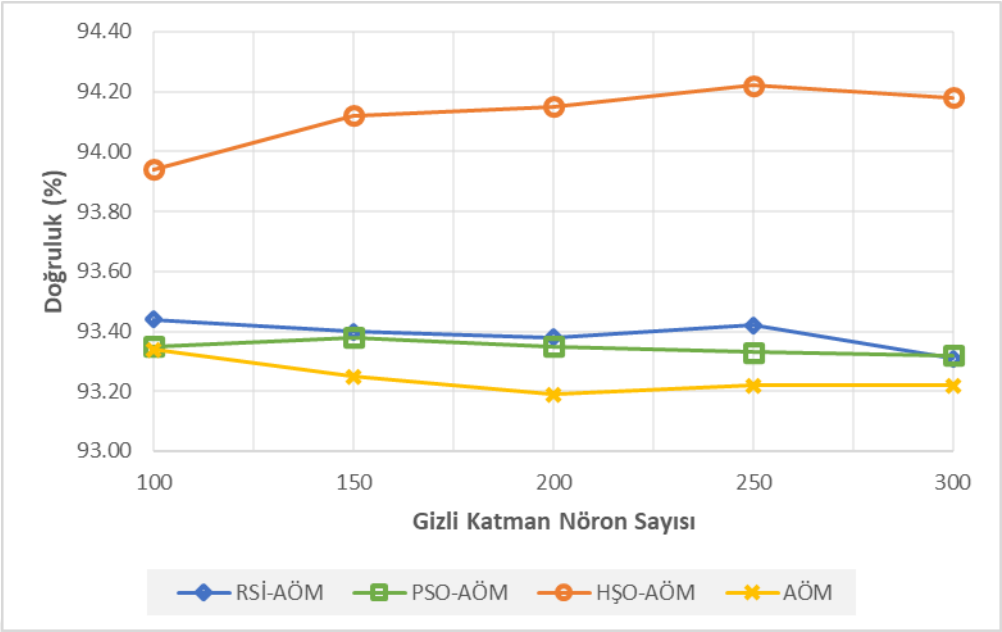
Gizli katmanda bulunan nöron sayısı olarak baktığımızda temel AÖM modelinin gizli katmanda bulunan 15 nöron sayısı ile diğer optimizasyon modellerine göre daha az nöron sayısında en iyi ortalama doğruluk değerini aldığı görülmektedir.

6.2.3. Tiroit Veri Seti için Sonuçlar

Şekil 6.13. ve Şekil 6.14.'te önerilen HŞO-AÖM modeli ile PSO-AÖM, RSİ-AÖM ve temel AÖM modellerinin test aşamasındaki ortalama yüzde doğruluk değerlerinin gizli katmandaki artan nöron sayısıyla değişim grafikleri verilmiştir. Şekil 6.13.'de görüldüğü üzere RSİ-AÖM ve PSO-AÖM optimizasyon modellerinin ortalama yüzde doğruluk değerleri birbirlerine yakın bir şekilde artarken, temel AÖM modeli görece daha yavaş bir artış sergilemiştir. HŞO-AÖM modeli ise artan nöron sayısı ile diğer tüm modellerin üstünde bir performans göstermiştir.



Şekil 6.13. Modellerin tiroit veri setinde [5-50] nöron aralığında test doğruluk oranı (%) değişimi



Şekil 6.14. Modellerin tiroit veri setinde [100-300] nöron aralığında test doğruluk oranı (%) değişimi

Şekil 6.14.'te görüldüğü üzere HŞO-AÖM modeli için diğer veri setlerinin aksine 100 nörondan sonra 250 nöron sayısına kadar test doğruluğu artmaya devam etmiştir. Diğer modeller ise hemen hemen aynı doğruluk değerlerini almaya devam etmişlerdir.

Tablo 6.11. Modellerin tiroit veri setinde gizli katmandaki nöron sayılarına bağlı olarak doğruluk değerlerinin maksimum, ortalama ve standart sapma değişimleri

M	HŞO-AÖM			RSİ-AÖM			PSO-AÖM			AÖM		
	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std	Maks. (%)	Ort. (%)	Std
5	93.22	92.97	.00136	92.58	92.58	.00000	92.63	92.59	.00015	92.58	92.58	.00000
10	93.71	93.47	.00115	93.13	93.07	.00049	93.04	92.99	.00031	92.61	92.54	.00062
15	93.63	93.41	.00110	93.28	93.24	.00026	93.22	93.19	.00027	92.64	92.61	.00022
20	93.60	93.45	.00125	93.32	93.29	.00018	93.31	93.27	.00020	92.92	92.76	.00078
25	93.72	93.52	.00141	93.32	93.30	.00018	93.32	93.27	.00025	93.07	92.98	.00053
30	93.72	93.59	.00081	93.33	93.31	.00015	93.32	93.29	.00023	93.19	93.13	.00032
35	93.76	93.65	.00102	93.33	93.30	.00020	93.31	93.28	.00023	93.24	93.20	.00024
40	93.79	93.64	.00086	93.35	93.32	.00029	93.33	93.30	.00021	93.32	93.24	.00042
45	93.85	93.74	.00068	93.39	93.33	.00033	93.35	93.32	.00028	93.29	93.25	.00028
50	93.79	93.70	.00084	93.39	93.35	.00027	93.35	93.32	.00017	93.32	93.27	.00031
100	94.03	93.94	.00065	93.53	**93.44**	.00053	93.42	93.35	.00051	93.39	**93.34**	.00037
150	94.35	94.12	.00094	93.46	93.40	.00061	93.50	**93.38**	.00071	**93.40**	93.25	.00075
200	94.28	94.15	.00108	93.51	93.38	.00082	93.50	93.35	.00106	93.28	93.19	.00061
250	**94.36**	**94.22**	.00082	**93.56**	93.42	.00118	**93.57**	93.33	.00086	93.39	93.20	.00129
300	93.22	92.97	.00136	93.49	93.31	.00096	93.46	93.32	.00116	93.33	93.20	.00106

Tablo 6.11.'de tiroit veri seti için gizli katmandaki nöron sayısına bağlı olarak modellerin toplam 50 çalıştırmada almış olduğu maksimum yüzde doğruluk değerleri, ortalama yüzde doğruluk değerleri ve doğruluk değerlerinin standart sapması (Std.) verilmiştir. Optimizasyon modelleri arasında önerilen HŞO-AÖM modeli gizli katmanda bulunan 250 nöronla 0.00082 standart sapma ile maksimum %94.36 doğruluk değerine, RSİ-AÖM modeli gizli katmanda bulunan 20 nöronla 0.01442 standart sapma ile maksimum %98 doğruluk değerine, ve PSO-AÖM modeli gizli katmanda bulunan 10 nöronla 0.01489 standart sapma ile maksimum %98 doğruluk değerine ulaşırken, temel AÖM modeli ise gizli katmanda bulunan 15 nöronla 0.01326 standart sapma ile maksimum %98.67 doğruluk değerine ulaşmıştır.

Tablo 6.12. Tiroit veri seti için modellerin çalışmasından alınan sonuçlar

Birim		**HŞO-AÖM**	**RSİ-AÖM**	**PSO-AÖM**	**AÖM**
Test Doğruluk Oranı	Ortalama (%)	**94.22**	93.44	93.38	93.34
	En İyi (%)	94.36	93.53	93.50	93.39
	En Kötü (%)	94.10	93.39	93.26	93.26
	Std	.00082	.00065	.00071	.00037
Gizli Katman Nöron Sayısı		250	100	150	100
Eğitim Süresi (sn)		1049.765	95.4757	375.5487	**0.1109**

Tablo 6.12.'de tiroit veri seti için modellerin toplamda 50 çalışma sonucunda en iyi ortalama yüzde doğruluk değeri veren gizli nörona ait sonuçlar verilmiştir. Standart sapma Std. olarak belirtilmiştir. Tabloyu incelediğimizde tiroit veri setinde tüm modellerin başarımının birbirine yakın olduğu görülmektedir. HŞO-AÖM modeli 0.00082 standart sapma ve %94.22 ortalama doğruluk değeri ile az farkla diğer modellerden daha iyi bir başarım elde etmiştir. RSİ-AÖM modeli 0.00065 standart sapma ve %93.44 ortalama doğruluk değeri ile sıralamada ikinci, PSO-AÖM modeli ise 0.00071 standart sapma ve %93.38 ortalama doğruluk değeri ile üçüncü gelmiştir. Karşılaştırması yapılan dört model arasından en düşük başarıma sahip model 0.00037 standart sapma ve %93.34 ortalama doğruluk değeri ile temel AÖM modeli olmuştur. Maksimum doğruluk oranı olarak baktığımızda %94.36 değeriyle HŞO-AÖM modeli üstün gelmiştir. Minimum test doğruluk oranını ise %93.26 değeri ile PSO-AÖM ve AÖM modelleri vermiştir. Gizli katmanda bulunan nöron sayısı olarak baktığımızda temel AÖM ve RSİ-AÖM modellerinin gizli katmanda bulunan 100 nöron sayısı ile diğer modellere göre daha az nöron sayısında en iyi ortalama doğruluk değerini aldığı görülmektedir.

7. SONUÇLAR VE ÖNERİLER

Önerilen optimizasyon modelinin doğruluk oranı başarımı literatürde AÖM optimizasyonu ile yapılan diğer çalışmalarla karşılaştırılmıştır. Tablo 7.1. incelendiğinde HŞO-AÖM modelinin hem iki sınıflı hem de çok sınıflı veri setlerinde AÖM parametre optimizasyonu alanında literatürdeki çalışmalara yakın sonuçlar verdiği görülmektedir.

Literatürde 7200 örnek bulunan 21 girişe sahip Tiroit veri setinde yapılan diğer optimizasyon çalışmalarının sonuçlarına bakıldığında %94.22 ortalama doğruluk değeri ile HŞO-AÖM modelinin üstün sonuç verdiği görülmektedir. İris veri setinde ise %97.13 ortalama doğruluk değeri ile HŞO-AÖM modelinin diğer çalışmalarla ortalama bir sonuç aldığı söylenebilir.

Tablo 7.1. HŞO-AÖM modelinin literatürde bulunan diğer optimizasyon algoritmaları ile oluşturulan modeller ile karşılaştırılması

Veri Seti	Referans	Algoritma	Başarı (%)	Önerilen (HŞO-AÖM)
Tiroit	(Berglund ve Belciug 2018)	KKO-AÖM	92.59	**%94.22**
	(Shukla ve ark 2016)	CB-AÖM	93.47	
İris	(Cao ve ark 2012)	KUE-AÖM	97.20	**%97.13**
	(Alçin 2015)	İGT-AÖM	98.00	
	(Pacifico ve Ludermir 2013)	CPSOS-AÖM	96.90	
	(Silva ve ark 2011)	GSO-AÖM	97.20	
Diyabet	(Zhu ve ark 2005)	E-AÖM	80.69	**%77.07**
	(Cao ve ark 2012)	KUE-AÖM	79.55	
	(Mohapatra ve ark 2015)	İGK-AÖM	78.50	
	(Alçin 2015)	İGT-AÖM	74.73	
	(Eshtay ve ark 2018)	RSİ-AÖM	77.15	
	(Pacifico ve Ludermir 2013)	CPSOS-AÖM	77.06	
Göğüs Kanseri	(Mohapatra ve ark 2015)	İGK-AÖM	97.77	**%96.68**
	(Alçin 2015)	İGT-AÖM	92.35	
	(Luo ve ark 2014)	SB-AÖM	97.22	
	(Silva ve ark 2011)	GSO-AÖM	96.90	
Şarap	(Cao ve ark 2012)	KUE-AÖM	97.95	**%98.31**
	(Pacifico ve Ludermir 2013)	CPSOS-AÖM	98.12	

Diyabet veri setinde AÖM optimizasyonu alanında yapılmış birçok araştırma bulunmaktadır. HŞO-AÖM modeli %77.07 ortalama doğruluk değeri ile bu veri setinde diğer çalışmalarla yaklaşık aynı başarımı elde etmiştir. Göğüs Kanseri veri setinde %96.68 ortalama doğruluk değeri ile literatürdeki İGT-AÖM modelinden daha iyi bir sonuç elde edilirken İGK-AÖM ve SB-AÖM modellerine kıyasla az farkla görece kötü bir sonuç elde edilmiştir. Şarap veri setinde ise %98.31 ortalama doğruluk değeri ile literatürdeki diğer çalışmalara yakın bir sonuç elde edilmiştir.

Bu tez çalışmasında diğer metasezgisel optimizasyon algoritmalarına göre henüz yeni tanıtılmış olan Harris Şahinleri Optimizasyonu (HŞO) algoritmasının AÖM'de rastgele verilen hiper parametrelerin iyileştirilmesindeki başarımı test edilmiştir. HŞO, 2019 yılında yayınlanan, sürü zekasına dayalı meta sezgisel optimizasyon algoritmasıdır. Üstün avlanma yeteneği olan Harris Şahinlerinin grup halinde avın izlemesi, tuzak kurma ve yakalama süreçlerinden esinlenilmiştir.

Önerilen HŞO-AÖM algoritmasının en önemli üstünlüğü AÖM için otomatik olarak en uygun giriş ağırlık değerlerini araması ve bulunmasıdır. Önemli bir dezavantajı ise gizli katmandaki nöron sayısının deneme yanılma metoduyla bulunmaya çalışılmasıdır. Bu da zaman alan bir yöntemdir. Aynı zamanda arama uzayının alt ve üst limitlerinin iyi seçilmesi algoritmanın yakınsama performansı için önemlidir.

Tüm veri setinde kullanılan 50 popülasyon sayısı ve 100 iterasyon sayısı ile 5000 çözüm kümesi içinden AÖM modeli için optimize edilmiş değerler bulunmuştur. İterasyon sayısının seçiminde literatürde yapılan diğer çalışmalar dikkate alınmıştır. Deneme yanılma metodu ile de modeller için ideal popülasyon boyutu bulunmuştur.

HŞO-AÖM modeli karşılaştırılan diğer modellere kıyasla Diyabet veri setinde elde ettiği %77.09 ortalama doğruluk değeri, İris veri setinde elde ettiği %97.13 ortalama doğruluk değeri ve Tiroit veri setinde elde ettiği %94.22 ortalama doğruluk değeri ile üstün gelmiştir. RSİ-AÖM modelinin tüm veri setlerinde diğer modellere kıyasla eğitim hızının daha düşük olması dikkat çekmektedir. Önerilen modelin dezavantajından bahsedecek olursak, RSİ-AÖM ve PSO-AÖM modellerine göre eğitimde daha fazla zaman almaktadır. Optimizasyon algoritmasında yapılacak iyileştirmelerle eğitim hızının minimize edileceği düşünülmektedir.

Alınan bu sonuçlarla sürü zekâsı algoritmalarından olan HŞO'nun temel AÖM'nin rastgele verilen gizli eşik değeri ve giriş ağırlıklarının optimal değerini bulunmasında etkili bir metot olduğu gösterilmiştir. Ayrıca AÖM'nin optimizasyonu ile sınıflandırmadaki başarım artırılmakla kalmamış, gizli katmanda bulunan nöron sayısı

da azalmıştır. Sunulan bu modelin sınıflandırma kabiliyetinin ölçülmesi için literatürde daha önce tanıtılmış olan PSO-AÖM ve RSİ-AÖM modelleriyle karşılaştırması yapılmıştır. Karşılaştırmanın adil olması açısından farklı performans ölçütleri ile sonuçlar verilmiştir. Genelleştirme performansı açısından baktığımızda HŞO-AÖM, PSO-AÖM, RSİ-AÖM ve AÖM'nin kullanılan beş veri setindeki üstün performanslarıyla diğer sınıflandırma algoritmalarına göre iyi birer sınıflandırıcı oldukları yapılan bu çalışmayla kanıtlanmıştır.

Çalışmada önerilen modelin optimizasyonu için HŞO algoritmasının orijinal versiyonu baz alınmıştır. Optimizasyonda herhangi bir iyileştirme yapılmamıştır. Bu sebeple karşılaştırılan diğer modellerin optimizasyonu için kullanılan PSO ve RSİ algoritmalarının aslına sadık kalınmıştır. Optimize edilen standart, bir diğer deyişle temel AÖM yönteminin 2006'da sunulan orijinal versiyonu korunmuştur.

Optimizasyonda yapılacak iyileştirmelerle ve AÖM'nin literatürde kabul görmüş geliştirilmiş versiyonlarının kullanılması ile sınıflandırma başarımı daha yüksek seviyelere çıkarılabilir aynı zamanda eğitimde geçen süre daha da aşağı çekilebilir.

KAYNAKLAR

Agatonovic-Kustrin S, Beresford R, 2000. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis. 22: 717-27.

Ahila R, Sadasivam V, Manimala K, 2015. An integrated PSO for parameter determination and feature selection of ELM and its application in classification of power system disturbances. Applied Soft Computing Journal, 32, 23-37.

Alçin ÖF, 2015. Aşırı Öğrenme Makinalarının Seyrek Geri Çatma Algoritmaları İle Optimizasyonu, Fırat Üniversitesi.

Altan G, 2018. Enhancing deep extreme learning machines: Novel multi-kernel autoencoders and implementation for detecting the chronic obstructive pulmonary disease, Doktora Tezi, Mustafa Kemal Üniversitesi, Fen Bilimleri Enstitüsü.

Berglund R, Belciug S, 2018. Improving extreme learning machine performance using ant colony optimization feature selection. Application to automated medical diagnosis. Annals of the University of Craiova-Mathematics and Computer Science Series, 45, 1, 151-5.

Bjerhammar A, 1951. Application of calculus of matrices to method of least squares: with special reference to geodetic calculations, Elander Göteborg, p.

Cao J, Lin Z, Huang GB, 2012. Self-adaptive evolutionary extreme learning machine. Neural Processing Letters, 36, 285-305.

Çatak FÖ, Mustacoglu AF, 2018. CPP-ELM: Cryptographically privacy-preserving extreme learning machine for cloud systems. International Journal of Computational Intelligence Systems.

Cheng R, Jin Y, 2015. A competitive swarm optimizer for large scale optimization. IEEE Transactions on Cybernetics, 45, 191-204.

de Oliveira JFL, Ludermir TB. An evolutionary extreme learning machine based on fuzzy fish swarms. Proceedings on the international conference on artificial intelligence (ICAI), 1.

Deng W, Zheng Q, Chen L, 2009. Real-Time Collaborative Filtering Using Extreme Learning Machine. 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, IEEE: 466-73.

Ding S, Xu X, Nie R, 2014. Extreme learning machine and its applications. Neural Computing and Applications, Springer. 25: 549-56.

Duan M, Li K, Yang C, Li K, 2018. A hybrid deep learning CNN–ELM for age and gender classification. Neurocomputing, 275, 448-61.

Eltaeib T, Mahmood A, 2018. Differential evolution: A survey and analysis. Applied Sciences (Switzerland).

Eshtay M, Faris H, Obeid N, 2018. Improving Extreme Learning Machine by Competitive Swarm Optimization and its application for medical diagnosis problems. Expert Systems with Applications, 104, 134-52.

Fisher RA, 1988, Iris Data Set [online]. https://archive.ics.uci.edu/ml/datasets/Iris [Ziyaret Tarihi: 21 Mayıs 2020].

Ghosh R, Verma B, 2003. A hierarchical method for finding optimal architecture and weights using evolutionary least square based learning. International Journal of Neural Systems, 13, 01, 13-24.

Gu S, Cheng R, Jin Y, 2018. Feature selection for high-dimensional classification using a competitive swarm optimizer. Soft Computing, 22, 811-22.

Han F, Yao HF, Ling QH, 2013. An improved evolutionary extreme learning machine based on particle swarm optimization. Neurocomputing, 116, 87-93.

Hassan R, Cohanim B, De Weck O, Venter G, (2005). A comparison of particle swarm optimization and the genetic algorithm. Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference.

Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja M, Chen H, 2019. Harris hawks optimization: Algorithm and applications. Future Generation Computer Systems.

Hornik K, 1991. Approximation capabilities of multilayer feedforward networks. Neural Networks.

Huang G-B, Babri HA, 1998. Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions. IEEE transactions on neural networks, 9, 1, 224-9.

Huang G-B, Zhou H, Ding X, Zhang R, 2011. Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 42, 2, 513-29.

Huang G-BB, Zhu Q-YY, Siew C-KK, 2006. Extreme learning Machine: Theory and Applications. Neurocomputing, 70, 489-501.

Huang G, Huang GB, Song S, You K, 2015. Trends in extreme learning machines: A review. Neural Networks.

Huang GB, Chen L, 2007. Convex incremental extreme learning machine. Neurocomputing, 70, 3056-62.

Huang GB, Chen L, Siew CK, 2006. Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Transactions on Neural Networks, 17, 879-92.

Ito Y, 1992. Approximation of continuous functions on Rd by linear combinations of shifted rotations of a sigmoid function with and without scaling. Neural Networks.

Kennedy J, Eberhart R, 1995. Particle swarm optimization. IEEE International Conference on Neural Networks - Conference Proceedings, IEEE. 4: 1942-8.

Lan Y, Soh YC, Huang G-B, 2009. Ensemble of online sequential extreme learning machine. Neurocomputing, 72, 13-15, 3391-5.

Leshno M, Lin VY, Pinkus A, Schocken S, 1993. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural networks, 6, 6, 861-7.

Liang N-Y, Huang G-B, Saratchandran P, Sundararajan N, 2006. A fast and accurate online sequential learning algorithm for feedforward networks. IEEE Transactions on neural networks, 17, 6, 1411-23.

Luo J, Vong CM, Wong PK, 2014, Sparse bayesian extreme learning machine for multi-classification [online], IEEE Transactions on Neural Networks and Learning Systems, 25, 836-43.

M. Forina SL, C. Armanino, 1991, Wine Data Set [online], https://archive.ics.uci.edu/ml/datasets/wine [Ziyaret Tarihi: 21 Mayıs 2020].

Matias T, Araújo R, Antunes CH, Gabriel D, 2013. Genetically optimized extreme learning machine. IEEE International Conference on Emerging Technologies and Factory Automation, ETFA.

Miche Y, Sorjamaa A, Bas P, Simula O, Jutten C, Lendasse A, 2010. OP-ELM: Optimally pruned extreme learning machine. IEEE Transactions on Neural Networks, 21, 158-62.

Moayedi H, Osouli A, Nguyen H, Rashid ASA, 2019. A novel Harris hawks' optimization and k-fold cross-validation predicting slope stability. Engineering with Computers, 1-11.

Mohapatra P, Chakravarty S, Dash PK, 2015. An improved cuckoo search based extreme learning machine for medical data classification. Swarm and Evolutionary Computation, 24, 25-49.

Moore EH, 1920. On the reciprocal of the general algebraic matrix. Bull. Am. Math. Soc., 26, 394-5.

Niu P, Ma Y, Li M, Yan S, Li G, 2016. A kind of parameters self-adjusting extreme learning machine. Neural Processing Letters, 44, 3, 813-30.

Pacifico LD, Ludermir TB. Evolutionary extreme learning machine based on particle swarm optimization and clustering strategies. The 2013 International Joint Conference on Neural Networks (IJCNN), 1-6.

Penrose R. A generalized inverse for matrices. Mathematical proceedings of the Cambridge philosophical society, 406-13.

Penrose R. On best approximate solutions of linear matrix equations. Mathematical Proceedings of the Cambridge Philosophical Society, 17-9.

Pouzols FM, Lendasse A, 2010. Evolving fuzzy optimally pruned extreme learning machine for regression problems. Evolving Systems, 1, 43-58.

Quinlan R, 1987, Thyroid Disease Data Set [online]. https://archive.ics.uci.edu/ml/datasets/Thyroid+Disease [Ziyaret Tarihi: 21 Mayıs 2020].

Raghuwanshi BS, Shukla S, 2018. Class-specific extreme learning machine for handling binary class imbalance problem. Neural Networks, 105, 206-17.

Sánchez-Monedero J, Hervas-Martinez C, Gutiérrez PA, Ruz MC, Moreno MCRi, Cruz-Ramirez M, 2010. Evaluating the performance of evolutionary extreme learning machines by a combination of sensitivity and accuracy measures. Neural Network World, 20, 899.

Saporetti CM, Duarte GR, Fonseca TL, Da Fonseca LG, Pereira E, 2018. Extreme Learning Machine combined with a Differential Evolution algorithm for lithology identification. Revista de Informática Teórica e Aplicada.

Saraswathi S, Sundaram S, Sundararajan N, Zimmermann M, Nilsen-Hamilton M, 2011. ICGA-PSO-ELM Approach for Accurate Multiclass Cancer Classification Resulting in Reduced Gene Sets in Which Genes Encoding Secreted Proteins Are Highly Represented. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 8, 452-63.

Shukla S, Yadav R, Naktode L. Correlation Based Extreme Learning Machine. 2016 9th International Conference on Developments in eSystems Engineering (DeSE), 268-72.

Silva DN, Pacifico LD, Ludermir TB. An evolutionary extreme learning machine based on group search optimization. 2011 IEEE Congress of Evolutionary Computation (CEC), 574-80.

Smith JW, Everhart, J.E., Dickson, W.C., Knowler, W.C., & Johannes, R.S., (1988), Pima Indians Diabetes Database [online], https://www.kaggle.com/uciml/pima-indians-diabetes-database [Ziyaret Tarihi: 21 Mayıs 2020].

Sun W, Wang C, Zhang C, 2017. Factor analysis and forecasting of CO2 emissions in Hebei, using extreme learning machine based on particle swarm optimization. Journal of Cleaner Production, 162, 1095-101.

Svozil D, Kvasnička V, Pospíchal J, 1997. Introduction to multi-layer feed-forward neural networks. Chemometrics and Intelligent Laboratory Systems.

Varol G, 2015. Aşırı öğrenme makineleri ile videolarda etkin hareket ve etkinlik tanıma, Yüksek Lisans Tezi, Boğaziçi Üniversitesi Fen Bilimleri Enstitüsü.

Wolberg WH, 1992, Breast Cancer Wisconsin (Original) Data Set [online]. https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(original) [Ziyaret Tarihi: 21 Mayıs 2020].

Yeom C-U, Kwak K-C, 2017. Short-Term Electricity-Load Forecasting Using a TSK-Based Extreme Learning Machine with Knowledge Representation. Energies, 10, 1613.

Zhang G, Eddy Patuwo B, Y. Hu M, 1998. Forecasting with artificial neural networks:: The state of the art. International Journal of Forecasting, 14, 35-62.

Zhang Y, Cai Z, Wu J, Wang X, Liu X. A memetic algorithm based extreme learning machine for classification. 2015 International Joint Conference on Neural Networks (IJCNN), 1-8.

Zhu QY, Qin AK, Suganthan PN, Huang GB, 2005. Evolutionary extreme learning machine. Pattern Recognition, 38, 1759-63.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı	**:**	Musa DOĞAN
Uyruğu	**:**	T.C.
Doğum Yeri ve Tarihi	**:**	Selçuklu / 13.06.1996
Telefon	**:**	05063381040
Faks	**:**	-
E-mail	**:**	musa.dogan@selcuk.edu.tr

EĞİTİM

Derece		**Adı, İlçe, İl**	**Bitirme Yılı**
Lise	:	Konya Lisesi (Anadolu), Meram, Konya	2014
Üniversite	:	Selçuk Üniversitesi, Selçuklu, Konya	2018
Yüksek Lisans	:	Selçuk Üniversitesi, Selçuklu, Konya	-

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**
2018-2019	Sentos Yazılım	Yazılım Geliştirici
2019-…	Selçuk Üniversitesi	Araştırma Görevlisi

UZMANLIK ALANI

Yazılım Geliştirme, Makine Öğrenmesi

YABANCI DİLLER

İngilizce

YAYINLAR

Dogan, M., & Ozkan, I.A. (2019). EEG-based Eye State Detection Using Artificial Neural Network. International Conference on Advanced Technologies (ICAT'19). 26-29 August, 2019, ISBN: 978-605-68537-4-6, Sarajevo, Bosnia and Herzegovina.

Dogan, M., & Ozkan, I.A. (2020). Sürü Zekası Yöntemleriyle Aşırı Öğrenme Makinesi'nin Öğrenme Parametreleri Optimizasyonu. Selçuk-Teknik Dergisi Cilt 19, Sayı:2-2020 **(Yüksek Lisans tezinden yapılmıştır.)**