



T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

SINIFLANDIRMA PROBLEMLERİNDE META-SEZGİSEL OPTİMİZASYON YÖNTEMLERİNİN ÖZELLİK SEÇİMİ VE AYRIKLAŞTIRMA AMACIYLA KULLANIMI

İsmail KOÇ

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Ocak - 2016

KONYA

Her Hakkı Saklıdır

ÖZET

YÜKSEK LİSANS TEZİ

SINIFLANDIRMA PROBLEMLERİNDE META-SEZGİSEL OPTİMİZASYON YÖNTEMLERİNİN ÖZELLİK SEÇİMİ VE AYRIKLAŞTIRMA AMACIYLA KULLANIMI

İsmail KOÇ

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Yrd. Doç. Dr. İsmail BABAOĞLU

2016, 181 Sayfa

Jüri
Yrd. Doç. Dr. İsmail BABAOĞLU
Doç. Dr. Erkan ÜLKER
Yrd. Doç. Dr. Onur İNAN

Sınıf etiketleri yardımıyla belli bir veri kümesi üzerinden oluşturulan modeller kullanılarak yeni örneklerin hangi sınıfa ait olacağının tahmin edilmesi genel olarak sınıflandırma problemi olarak adlandırılmaktadır. Birçok alanda karşılaşılan bu problemlerin çözümü için farklı disiplinlerde yeni yöntemler üzerine çalışmalar yapılmaktadır. Dolayısıyla her geçen gün yeni yaklaşımlar sunulmakta ve çözüm yöntemleri geliştirilmektedir. Bununla birlikte sınıflandırma başarısının artırılması amacıyla da literatürde farklı teknikler yer almaktadır. Bu tezde veri madenciliğinde kullanılan ve önemli yöntemlerden biri olan özellik seçimi kullanılmıştır. Alt küme seçimi olarak bilinen özellik seçimi makine öğrenmesinde yaygın olarak kullanılan bir yöntemdir. Bu yöntem, veriyi işleme ve analiz etmek için yönetilebilir boyuttaki veriyi azaltan teknikleri ve araçları tanımlayan bir terimdir. Özellik seçimi işleminde, veri kümesinden elde edilen özellik alt kümesi öğrenme algoritması uygulaması için seçilir. En iyi alt küme, çözüm uzayı için en yüksek doğruluk oranına sahip olan en küçük boyutlu veri kümesinden oluşur. Veri kümesindeki geriye kalan önemsiz nitelikler ise yok sayılır. Bu işlem, önemli bir veri ön işleme aşamasıdır.

Problemlerde karşılaşılan veriler sürekli veya kesikli (ayrık) veri şeklinde olabilmektedir. Özellikle tahmin modelleri oluşturma çalışmalarında kesikli veri tercih edilmektedir. Bu tercihin sebebi ise ayrık verilerin bilgi düzeyli gösterilebilir olması, bazı işlemler sonrası sadeleştirilmiş olması, anlaşılır ve açıklanabilir olmasıdır. Sürekli verinin kesikli veriye dönüştürülmesi işlemleri genel olarak "veri ayrıklaştırma" olarak tanımlanmaktadır. Ayrıklaştırmanın başarısı hangi algoritmanın kullanıldığına, verinin dağılımına ve sonuç çıkarma modeli gibi parametrelere bağlıdır.

Bu tezde optimizasyon algoritmalarının özellik seçimi ve ayrıklaştırma amacıyla kullanılması araştırılmıştır. Dört farklı global erişilebilir veri kümesi üzerinde özellik seçimi, eşit genişlik ve eşit frekansa göre ayrıklaştırma amacıyla Yapay Arı Kolonisi, Guguk Kuşu, Yarasa ve Yerçekimsel Arama algoritmaları kullanılmış olup analiz sonuçları karşılaştırmalı olarak sunulmuştur. Süre analizleri, özellik seçimi sonrası elde edilen özellikler ve ayrıklaştırma sonrası belirlenen ayrıklaştırma sınırlarına ait analizler ise tez kapsamı dışında tutulmuştur. Özellik seçimi işlemlerinde optimizasyon algoritmalarının ikili versiyonu kullanılmış olup ayrıklaştırma işlemlerinde ise algoritmaların sürekli versiyonları kullanılmıştır.

Anahtar Kelimeler: Ayrıklaştırma, Destek Vektör Makinaları , Eşit Genişlik, Eşit Frekans, Guguk Kuşu Arama Algoritması, İkili Optimizasyon Algoritmaları, Optimizasyon Algoritmaları, Özellik Seçimi, Yapay Arı Kolonisi Algoritması, Yarasa Algoritması, Yerçekimi Arama Algoritması

ABSTRACT

MS THESIS

UTILIZATION OF METAHEURISTIC OPTIMIZATION METHODS FOR FEATURE SELECTION AND DISCRETIZATION ON CLASSIFICATION PROBLEMS

Ismail KOC

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

Advisor: Asst. Prof. Dr. Ismail BABAOGLU

2016, 181 Pages

Jury
Asst. Prof. Dr. Ismail BABAOGLU
Assoc. Prof. Dr. Erkan ULKER
Asst. Prof. Dr. Onur INAN

Prediction of the samples classes using models which are formed through a given data set by means of the class labels is generally named as classification problem. In order to solve these kinds of problems encountered in several areas, many researches on novel methods are studied in different disciplines. Therefore, novel approaches have been presented, and solution methods have been developed day by day. Besides, there are different techniques which are used for increasing the classification accuracy in literature. In the thesis, feature selection which is one of the important techniques used in data mining has been utilized. Feature selection known as subset selection is a method which is commonly used in machine learning. This method is a term which defines resources and techniques of decreasing data with manageable dimension for operation and analysis of data. The subset of the features which are obtained from the dataset is selected for the application of the learning algorithm in the feature selection process. The best subset consists of data set with the least dimensions that has the maximum accuracy. The remaining redundant attributes are disregarded. This process is one of the important data preprocessing stages.

The data encountered in problems can be in discrete or continuous data form. The discrete data is preferred in the studies in forming prediction models, especially. The reason of this preference can be said that discrete values can be shown as in information level, they are summarized in the end of some processes and they are understandable and explicable. The processes transforming continuous data into discrete data are generally described as data discretization. Discretization is a data preprocessing approach used frequently in methods of data mining and machine learning. The success of discretization process is related to the parameters such as the result attainment model, data distribution and which algorithm is utilized.

In the thesis, the usage of optimization algorithms in the purpose of feature selection and discretization has been studied. Artificial Bee Colony, Bat, Gravitational Search and Cuckoo Search Algorithms have been used with the intention of feature selection, equal width discretization, equal frequency discretization using four different global available data set and their analysis results have been presented comparatively. However, time analysis and the analysis of both the features obtained by feature

selection process and the boundaries obtained by the discretization process are excluded from this thesis. While the binary versions of these algorithms have been used in the processes of feature selection, the continuous versions of them have been utilized in the discretization processes.

Keywords: Artificial Bee Colony Algorithm, Bat Algorithm, Binary Optimization Algorithms, Cuckoo Search Algorithm, Discretization, Equal Frequency, Equal Width, Feature Selection, Gravitational Search Algorithm, Optimization Algorithms, Support Vector Machines

ÖNSÖZ

Bu çalışmamda beni destekleyip, yönlendiren danışman hocam Yrd. Doç. Dr. İsmail Babaoğlu'na ve Selçuk Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü Öğretim Eleman'larına teşekkürlerimi sunarım.

Çalışmalarım sırasında benden maddi ve manevi desteklerini asla esirgemeyen aileme en içten teşekkürlerimi sunarım.

İsmail KOÇ
KONYA-2016

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vii**

İÇİNDEKİLER ..... **viii**

SİMGELER VE KISALTMALAR ..... **x**

1. GİRİŞ **1**
1.1. Tez Çalışmasının Amacı 2
1.2. Tez Çalışmasının Önemi 4

2. KAYNAK ARAŞTIRMASI **5**
2.1. Optimizasyon Algoritmaları Yardımıyla Sınıflandırma ve Destek Vektör Makinaları ile İle İlgili Literatür Çalışmaları 5
2.2. Özellik Seçimi ile İle İlgili Literatür Çalışmaları 11
2.3. Ayrıklaştırma ile İle İlgili Literatür Çalışmaları 18

3. OPTİMİZASYON ALGORİTMALARI **21**
3.1. Yapay Arı Koloni Algoritması 21
3.1.1. Arıların yem bulma davranışları 21
3.1.3. Yiyecek kaynakları 22
3.1.3. Görevi belirli işçi arılar 22
3.1.4. Görevi belirli olmayan işçi arılar 22
3.1.5. Yapay arı kolonisi algoritması ve temel adımları 24
3.1.6. İkili (Binary) yapay arı koloni algoritması 25
3.2. Yerçekimsel Arama Algoritması 28
3.2.1. İkili (binary) yerçekimsel arama algoritması 31
3.3. Guguk Kuşu Arama Algoritması 32
3.3.1. Başlangıç popülasyonu oluşum evresi 35
3.3.2. Guguk kuşunun yumurtlama evresi 36
3.3.4. En kötü habitattaki guguk kuşlarını yok etme evresi 38
3.3.5. Yakınsama evresi 39
3.3.6. İkili guguk kuşu arama algoritması 39
3.4. Yarasa Algoritması 40
3.4.1. İkili Yarasa Algoritması 43

4. DESTEK VEKTÖR MAKİNALARI **44**
4.1. Çekirdek (Kernel) Fonksiyonları 45

5. ÖZELLİK SEÇİMİ VE ÖNERİLEN YÖNTEM **47**

5.1. Özellik Seçimi 47
5.1.1. Özellik Seçimi Tanımları 47
5.2. Özellik Seçiminin Avantajları ve Dezavantajları 48
5.3.1. Arama organizasyonu 49
5.3.2 Sonraki jenerasyon 50
5.4. Özellik Seçimi İçin Önerilen Sistem 50

6. VERİ AYRIKLAŞTIRMA TEKNİKLERİ VE ÖNERİLEN YÖNTEM 53
6.1. Veri Ayrıklaştırma İşlemi 53
6.2. Karakteristik Bir Ayrıklaştırma İşlemi 53
6.2.1. Kesim noktası seçme 54
6.2.3. Bölme (split) ve birleştirme (merge) 55
6.2.4. Durdurma kriteri 55
6.3. Veri Ayrıklaştırma Yöntemlerinin Sınıflandırılması 55
6.4. Veri Ayrıklaştırmada Kullanılan Bazı Yöntemler 57
6.4.1. Eşit genişlikli ayrıklaştırma 57
6.4.2. Eşit frekanslı ayrıklaştırma 58
6.5. Ayrıklaştırma İçin Önerilen Sistem 59

7. TEST VE DEĞERLENDİRME 63
7.1. Deneysel Çalışmalar 63
7.2. Tanı Testlerinde Kullanılan Ölçütler 64
7.3. Özellik Seçimi İçin Test Sonuçları 65
7.3.1. WPBC veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar 65
7.3.2. BCW veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar 68
7.3.3. QSAR veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar 69
7.3.4. DRD veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar 71
7.4. Eşit Genişlikli Ayrıklaştırma Yöntemi İçin Test Sonuçları 74
7.4.1. WPBC verisi kullanılarak EWD yöntemine ait deneysel sonuçlar 74
7.4.2. BCW veri kümesi kullanılarak EWD yöntemine ait deneysel sonuçlar 75
7.4.3. QSAR veri kümesi kullanılarak EWD yöntemine ait deneysel sonuçlar 78
7.4.4. DRD veri kümesi kullanılarak EWD yöntemine ait deneysel sonuçlar 79
7.5. Eşit Frekanslı Ayrıklaştırma Yöntemi İçin Test Sonuçları 82
7.5.1. WPBC verisi kullanılarak EFD yöntemine ait deneysel sonuçlar 82
7.5.2. BCW verisi kullanılarak EFD yöntemine ait deneysel sonuçlar 84
7.5.3. QSAR veri kümesi kullanılarak EFD yöntemine ait deneysel sonuçlar 86
7.5.4. DRD veri kümesi kullanılarak EFD yöntemine ait deneysel sonuçlar 88
7.6. Uygulama Sonuçlarının Değerlendirilmesi 89

8. SONUÇLAR VE ÖNERiLER 97
8.1. Sonuçlar 97
8.2. Öneriler 97

KAYNAKLAR 99

EKLER 108

ÖZGEÇMİŞ 181

SİMGELER VE KISALTMALAR

Simgeler

v	: hız faktörü
x	: konum değeri
f	: frekans
γ	: Çekirdek Boyutu

Kısaltmalar

ABC	: Yapay Arı Kolonisi Algoritması
GSA	: Yerçekimsel Arama Algoritması
CSO	: Guguk Kuşu Arama Algoritması
BA	: Yarasa Algoritması
DVM	: Destek Vektör Makinaları
ÖS	: Özellik Seçimi
EWD	: Eşit Genişlikli Ayrıklaştırma
EFD	: Eşit Frekanslı Ayrıklaştırma
BCW	: Breast Cancer Wisconsin
DRD	: Diabetic Retinopathy Debrecen
QSAR	: QSAR biodegradation
WPBC	: Breast Cancer Wisconsin (Prognostic)
KKA	: Karınca Koloni Algoritması
YBS	: Yapay Bağışıklık Sistemi
ÖAS	: Özellik Alt Küme Seçimi
DVR	: Destek Vektör Regresyonu
KNN	: K-En Yakın Komşuluk
IİÇAO	: Isıl İşlem Tabanlı Çok Amaçlı Optimizasyonu
PSO	: Parçacık Sürü Optimizasyonu
GBYSA	: Geri Beslemeli Yapay Sinir Ağları
GA	: Genetik Algoritma
YSA	: Yapay Sinir Ağları
DEA	: Diferansiyel Gelişim Algoritması

1. GİRİŞ

Özellik seçimi (ÖS), sınıflandırma problemlerinde ve birçok öğrenme algoritmalarında önemli bir yer teşkil etmektedir. ÖS, sonuç değeriyle ilgili olan, en faydalı ve en önemli özelliklerin seçilerek veri kümesine ait özellik sayısının azaltılması için kullanılan bir yöntemdir. Amaç; hesaplama yükünü azaltıp başarı oranını yükseltmektir. Literatürde ÖS yöntemleri ile ilgili çok fazla araştırma bulunmaktadır ve birçok alanda çeşitli uygulamaları ile karşılaşılmaktadır. (Zongker ve Jain 1996, Onnia ve ark 2001);

1) Özellik sayısının çok olduğu durumda veri kümesindeki gizli ilişkileri elde etmek amacıyla,

2) Çok fazla sayıda algılayıcıdan veri elde edilen uygulamalar,

3) Sınıflandırma amacıyla bir araya getirilen çoklu modellere ait parametre sayısının çok olduğu durumlar,

4) Giriş uzay boyutunun azaltılması gereken uygulamalardır.

ÖS yöntemleri başlıca iki ana sınıfta incelenebilir. Bunlardan ilki istatistiğe dayalı yöntemlerdir. Diğeri ise sınıflandırmaya dayalı yöntemler olarak bilinmektedir (Onnia ve ark 2001). İstatistiğe dayalı yöntemler sınıflandırma problemine bağlı kalmadan istatistiksel yöntemleri kullanarak en uygun özellik alt kümesini tespit etmeyi amaçlamaktadır. Sınıflandırmaya dayalı yöntemler ise sınıflandırma probleminde kullanılmayacak olan özellik alt kümelerinin tespit edilip veri kümesinden atılmasına veya sınıflandırma için en önemli özelliklerin bulunmasına dayalı yöntemler olarak bilinir.

ÖS makine öğrenmesinde yaygın olarak kullanılan bir yöntem olup literatürde alt küme seçimi olarak bilinmektedir. ÖS işleminde, veri kümesinden elde edilen özellik alt kümesi, öğrenme algoritması için seçilir. Çözüm uzayı için en yüksek doğruluk oranına sahip olan en küçük boyutlu veri kümesinden oluşan küme en iyi alt küme olarak kabul edilir. Veri kümesindeki geriye kalan önemsiz özellikler yok sayılır. Bu aşama, önemli bir veri ön işleme aşamasıdır. ÖS'nin temel hedefi, orijinal özelliklerin hepsini kullanmadan en yüksek oranda veri bütünlüğünü sağlamaktır. Bununla birlikte minimum özellik alt kümesini bulmaktır. Birçok gerçek dünya probleminde, gereksiz, yanıltıcı veya gürültülü verinin çokluğundan dolayı ÖS bir

zorunluluk olarak kabul edilmektedir. ÖS sonuçlarında en ideal çözümü bulmak için, tüm özellik alt kümelerinin test edilmesi gerekmektedir.

Veri ön işleme adımlarından birisi de veri ayrıklaştırma yöntemidir. Kısaca sürekli verinin kesikli veriye dönüştürülmesi olarak ifade edilmektedir. Buna ek olarak veri ayrıklaştırma için aynı anlamı taşıyan literatürde farklı tanımlar yer almaktadır; Jin ve ark (2007), veri ayrıklaştırma işlemini en az veri kaybı ile sürekli verinin sonlu komşu aralıklar şekline dönüştürülme evresi olarak tanımlarken, Das ve Vyas (2010) aynı veya yakın özelliklere sahip sürekli verinin gruplara veya aralıklara dönüştürülmesi süreci olarak ifade etmektedir. Veri ayrıklaştırma işlemine verilebilecek en basit örneklerden birisi ise, yaş niteliğine ait değerlerin veri ayrıklaştırma sonrası elde edilecek değer aralıklarına göre dağıtılıp bu aralıkların "genç", "orta yaş" ve "yaşlı" gibi farklı üç grup olarak ifade edilmesidir.

Veri ayrıklaştırma aşaması sonrasında sürekli veri azaltılıp veri kümesi en iyi temsil edilecek şekilde özetlenirken, veri madenciliği sonrası elde edilen bilgilerin kullanımı ve sunumu kolaylaşmış, bununla birlikte bu bilgilerin daha anlamlı hale gelmesi de sağlanmış olur (Chakrabarti ve ark 2008). Veri ayrıklaştırma yöntemi için birçok yöntem kullanılmakla beraber mevcut yöntemlerin yanı sıra yöntemler üzerine yeni özellikler de eklenerek farklı ayrıklaştırma yöntemleri geliştirilmiştir.

Bu çalışmanın ilk bölümünde tezin amacı ve önemi belirtilmiştir. Sonraki bölümlerde ise sırasıyla kaynak araştırması, optimizasyon algoritmaları, DVM, ÖS ve uygulanan yöntem, ayrıklaştırma ve önerilen yöntem, test ve değerlendirme ve sonuçlar ve öneriler anlatılmıştır. Geliştirilen ÖS ve eşit frekans ve eşit genişlik tabanlı ayrıklaştırma yöntemleri MATLAB2014 ortamında kodlanmıştır. Uygulamalar algoritmalara ait farklı parametre değerlerinde test edilecek şekilde tasarlanmıştır. Çalışmanın son bölümünde ise geliştirilen yöntemlerin uygulama sonuçları analiz edilmiştir.

1.1. Tez Çalışmasının Amacı

ÖS makine öğrenmesi ve istatistikte değişken seçimi, nitelik seçimi veya değişken alt küme seçimi olarak bilinmektedir. ÖS; veri modeli oluşturmada kullanılmak üzere veri kümesine ait özelliklerin bir alt kümesini seçme aşaması olarak tanımlanır. ÖS; örüntü tanımada, istatistikte ve veri madenciliği alanında yaygın olarak

tercih edilen bir yöntemdir. ÖS, sistemin çıktı olarak belirlediği konu ile bağlantılı olan, sınıflandırma sürecinde en faydalı ve en önemli özelliklerin seçilerek veri kümesine ait özellik sayısının indirgenmesi için kullanılmaktadır. Bu işlem sayesinde, hesaplama yükü azalırken, daha yüksek sınıflandırma başarısı hedeflenmektedir. Özellik seçim metotları üç ana sınıfta toplanmaktadır. Bunlar sırasıyla; filtreleme metodu, *Wrapper* metodu ve Embedded metodudur.

Uygulamalarda genellikle karşılaşılan veri tipleri sürekli veya kesikli (ayrık) veri biçiminde olmaktadır. Özellikle tahmin modelleri geliştirme çalışmalarında sürekli veriden daha çok kesikli veri tercih edilmektedir. Bu tercih sebepleri ise, ayrık verilerin bilgi düzeyli şekilde gösterilebilir olması, bazı işlemler sonucu sadeleştirilmiş olması, anlaşılır ve açıklanabilir olması şeklinde sayılabilir (Olson ve Delen 2008). Sürekli verilerin kesikli verilere dönüştürülmesi için uygulanan işlemler "veri ayrıklaştırma" olarak tanımlanmaktadır (Koçoğlu 2012). Ayrıklaştırma, veri madenciliği ve makine öğrenmesi yöntemlerinde çok sık kullanılan/karşılaşan bir veri önişleme yöntemidir (Xi ve Ouyang 2000, Hsu ve ark 2003). Sürekli verileri ayrıklaştırmak için herhangi bir ön bilgiye gerek duymadan ayrıklaştırma yöntemleri sürekli verileri birçok alana gruplamaktadır. Her bir bölge farklı isimlerle adlandırılmakta ve her bir veri kendi bölgesine ilişkin isimlere atanmaktadır. Ayrıklaştırmanın başarısını etkileyen farklı parametreler bulunmaktadır. Bu parametreler; kullanılan algoritma, verinin dağılımı ve sonuç çıkarma modeli şeklinde ifade edilebilir. Verilerin ayrık hale getirilmesinde kullanılan birçok teknik bulunmaktadır. Bu çalışmada, EWD (Hsu ve ark 2003), EFD (Jiang ve ark 2009) yöntemleri kullanılmıştır.

Bu tezde optimizasyon algoritmalarının ÖS amacıyla kullanılması araştırılmıştır. Bu algoritmalar; Yapay Arı Kolonisi Algoritması, Yer çekimi Arama Algoritması, Yarasa Algoritması ve Guguk Kuşu Arama Algoritmalarıdır. Bunun yanı sıra veri kümelerindeki sürekli veriler literatürde yer alan farklı ayrıklaştırma işlemlerine tabi tutularak veriler ayrık hale getirilmiş ve elde edilen ayrık veri kümesi orijinal veri kümesi ile başarısı kıyaslanmıştır. Farklı global erişilebilir veri kümesi üzerinde ÖS ve ayrıklaştırma amacıyla farklı optimizasyon algoritmaları kullanılmış ve analiz sonuçları karşılaştırmalı olarak sunulmuştur.

1.2. Tez Çalışmasının Önemi

Özellikle son birkaç yılda yüksek boyutlu verilerin büyük ölçüde arttığı gözlemlenmiştir. Veri boyutunun yüksek olmasından dolayı hesaplama maliyeti ve hesaplama süresi benzeri değerlerde yüksek artış görülmektedir. Bu durumlardan dolayı ÖS, veri madenciliğinde önemli bir yere sahiptir. Bu çalışma, ÖS yöntemi ile dört farklı global erişilebilir veri kümelerinde, veri kümesini temsil edecek en iyi özellik alt kümesinin bulunmasına katkı sağlayacaktır.

Gerçek dünya problemlerinde problemin veri uzayı genellikle sürekli verilerden oluşmaktadır. Sürekli verilerin anlaşılırlık ve açıklanma düzeyi daha karmaşık olduğundan ayrıklaştırma yöntemlerine ihtiyaç duyulmaktadır. Bu sebeplerden ötürü, ayrıklaştırma yöntemleri ile de sürekli veriler ayrık hale getirilerek bilgi düzeyli ve sade bir formata getirilecektir. Ayrık hale getirilen veri kümeleri üzerinde sınıflandırma işlemi sonucu elde edilen eğitim ve test başarıları, orijinal veri setleri ile kıyaslanarak ayrıklaştırmanın başarısı ve ne kadar önemli olduğu görülecektir.

2. KAYNAK ARAŞTIRMASI

2.1. Optimizasyon Algoritmaları Yardımıyla Sınıflandırma ve Destek Vektör Makinaları ile İle İlgili Literatür Çalışmaları

García Nieto ve ark (2016), *Spirulina platensis*'in başarılı bir büyüme evresini tahmin etmek amacıyla DVM'yi PSO ile birlikte kullanarak hibrit bir model önermiştir. Bu optimizasyon mekanizması DVM'nin eğitim aşamasındaki çekirdek parametre ayarlamasını içermektedir. Bu çalışmada DVM'nin doğrusal, kuadratik ve RBF olmak üzere üç farklı regresyonu yapılmıştır. Deneysel veriler PSO-DVM modelinin iyi bir model olduğunu doğrulamıştır.

Aydin ve ark (2011), DVM'nin parametre optimizasyonu için çok amaçlı yapay bağışıklık Algoritmasını kullanmışlardır. Önerilen sistemde DVM'nin çekirdek ve ceza parametrelerini optimize etmek hedeflenmiştir. Önerilen algoritma, anomali tespit problemleri ve indüksiyon motorlarının hata tespitine başarılı bir şekilde uygulanmış ve başarılı sonuçlar elde edilmiştir.

Cheng ve Juang (2011), TS tipli bulanık sınıflandırıcı ile eğitilmiş artımsal bir DVM sunmuşlardır (ISVM-FC). Önerilen bu yöntem Gaussian çekirdekli DVM ile kıyaslanmış ve test süresinin ve eğitim başarısının daha iyi olduğu tespit edilmiştir. Buna ek olarak önerilen yöntemin genelleştirme özelliğini kaybetmeden sınıflandırıcı depolaması için hafıza kullanımını da azalttığı görülmüştür.

Choi ve ark (2011), ret seçenekli DVM'yi kullanarak kanser sınıflandırması için gen tahmini ve seçimi yapmışlardır. Sonuçlar standart DVM ile kıyaslandığında önerilen DVM'nin tahminsel hataları ciddi oranda azalttığı görülmüştür.

Gang ve Zhuping (2011), DVM ile PSO'yu birlikte kullanarak trafik güvenliği tahmin modeline uygulamışlardır. GBYSA'nın dezavantajlarını çözmek için bu yaklaşımı önermişlerdir. 1970-2006 yılları arasında Çin'deki trafik verileri alınarak yapılan çalışmada PSO-DVM'nin GBYSA'dan daha iyi tahminler ürettiği görülmüştür.

Liu ve ark (2011), ÖS için iyileştirilmiş PSO'yu önermişlerdir. Çalışmada ayrık problemlerin çözümü için çoklu sürü PSO (ÇSPSO)'yu önermişler. Önerilen bu algoritma ile DVM birleştirilerek önerilen iyileştirilmiş ÖS'nin (İÖS) standart PSO ve GA'ya göre tahmin kesinliği açısından çok daha iyi sonuçlar ortaya koyduğu belirtilmiştir.

Rubio ve ark (2011), zaman serisi tahmini uygulamasını gerçekleştirmek için LS-DVM'de parametre seçimi için sezgisel bir yöntemi literatüre kazandırmışlardır.

Wang ve Meng (2011), DVM'nin parametre seçimi algoritması olarak PSO'yu kullanmışlar ve önerilen modelin YSA'dan hızlı yakınsama, yüksek genelleştirme yeteneği ve basit yapısı gibi özelliklerinden dolayı daha avantajlı olduğunu ifade etmişlerdir.

Yu ve ark (2011), görüntü bölütlemeye uygulamak için destek vektörlerinin özelliklerine ve budama stratejisine dayalı bir modifiye DVM önermişlerdir. Önerilen yaklaşım gerçek görüntüler üzerinde test edilmiş ve hesaplama maliyetinde önemli bir düşüş sağlanmıştır. Bununla birlikte önerilen DVM'nin görüntü bölütleme işleminde de oldukça iyi çalıştığı gözlemlenmiştir.

Bhadra ve ark (2012), meta sınıflandırıcı tasarımı için DEA tabanlı DVM parametrelerinin optimizasyonu üzerinde çalışmışlardır. Önerilen meta sınıflandırıcının klasik sınıflandırıcılardan daha iyi olduğu tespit edilmiştir.

Boubezoul ve Paris (2012) *Cross-Entropy* isimli metodun kullanımına dayalı global bir optimizasyon yaklaşımı ile eş zamanlı hem hiper parametre hem de özellik alt kümesinin optimizasyonunu gerektiren problemlerin çözümünü hedeflemişlerdir.

Dehuai ve ark (2012), Isıl İşlem Algoritması kullanarak DVM tabanlı bir sıcaklık tahmini yapmışlardır. Önerilen yöntemi PSO-DVM ile kıyasladıklarında yeni yöntemin daha iyi sonuç elde ettiği tespit edilmiştir.

Guo ve ark (2012), GA kullanarak iyileştirilmiş DVM tabanlı petrol fiyat tahmin modelini geliştirmişlerdir. Önerilen sistemde, DVM'nin ceza parametresi ve çekirdek parametresi GA ile optimize edilmiş olup gerçek dünya verileri kullanılarak elde edilen sonuçlar kıyaslandığında önerilen sistemin standart DVM'den daha iyi sonuç verdiği görülmektedir.

Hsieh ve ark (2012), PSO ve ABC Algoritması ile melezleştirilmiş sisteme dayanan ceza güdümlü DVM kullanılarak mali sıkıntı eğilim verileri üzerinde veri madenciliği çalışması yapmışlardır.

Huang (2012), DVM'yi GA ile birlikte kullanarak hibrit bir stok seçim modelini literatüre kazandırmıştır. Bu çalışmada, GA model parametrelerinin optimizasyonu için kullanılmıştır.

Nikitidis ve ark (2012), çok destekli DVM'lerin artımsal eğitimi için çarpımsal güncelleme kurallarını önermişlerdir. Önerilen çarpımsal kuralları herhangi bir öğrenme oranı parametresinden bağımsız olarak gerçekleştirmişlerdir.

Sahu ve Mishra (2012), kanser mikro dizi verileri için PSO kullanarak yeni bir özellik seçim algoritmasını (ÖSA) önermişlerdir. Önerilen yöntemi iki aşamalı olarak

gerçekleştirmişlerdir. İlk aşamada, k-means kümeleme kullanarak verileri kümele işlemini yapmışlar; ikinci evrede ise, yeni özellik alt kümelerini PSO'ya giriş değeri olarak vermişlerdir. Sonuçlar incelendiğinde, önerilen sistemin literatürdekilere göre daha iyi olduğu gözlemlenmiştir.

Sartakhti ve ark (2012), Isıl İşlem Algoritması ile birlikte DVM'ye dayanan yeni bir hibrit sistem kullanarak hepatit hastalıklarının teşhisi üzerinde çalışma yapmışlardır. UCI makine öğrenmesi veri tabanından alınan veri setleri üzerinde yapılan testlerde önerilen hibrit sistemin literatürdeki diğer sınıflandırma metotlarına göre çok daha iyi olduğu görülmüştür.

Zhou ve ark (2012), Karınca Koloni Algoritması (KKA) ile birlikte DVM regresyonu kullanarak kömür yakıtlı kazanlardan ortaya çıkan NO_x emisyonu modellemesi üzerinde bir çalışma yapmışlardır. Çalışmada kullanılan KKA'nın yüksek oranında DVM'nin optimal parametrelerini otomatik olarak elde ettiği görülmüştür.

Zhou ve ark (2012), sezgisel yöntemlerle DVM'yi birlikte kullanarak yer altındaki kaya patlamalarına uzun vadeli tahmin modeli geliştirmeyi önermişlerdir. Sezgisel yöntem olarak PSO ve GA'yı çalışmalarında kullanmış olan Zhou ve ark (2012) bu yöntemleri DVM ile hibrit olarak kullanmışlardır. PSO-DVM ve GA-DVM isimli geliştirdikleri hibrit modelleri DVM'nin grid arama yöntemiyle karşılaştırmışlardır. Ortaya çıkan sonuçlar yeni yöntemin ikisinin de grid arama yönteminden daha hızlı olduğunu göstermekle beraber kararlı bir yapıya sahip olduğunu da göstermiştir.

Bao ve ark (2013), DVM'lerin parametre optimizasyonu için PSO ve örüntü arama tabanlı memetik bir algoritma geliştirmişlerdir. Önerilen memetik algoritmasının kararlı ve etkili olduğu görülmüştür.

Che (2013), adaptif PSO (APSO) ve optimal eğitim alt kümesine dayanan DVM regresyonu üzerinde bir çalışma yapmıştır. Önerilen yöntemde DVM'nin parametre seçimi için APSO kullanılmıştır. UCI veri seti ve elektrik yük tahmin verileri kullanılarak elde edilen deneysel sonuçlar incelendiğinde önerilen yöntemin başarılı olduğu ve daha iyi genelleştirme performansı ürettiği tespit edilmiştir.

Chen ve ark (2013), DVM'leri hızlandırmak için hızlı ÖS yöntemini önermişlerdir. Bu çalışmada özellikle çok sınıflı problemler için ÖS yöntemini önermişlerdir. Çeşitli veri kümeleri üzerinde deneyler yapılmış ve önerilen yöntemin en rekabetçi algoritmalara göre çoğu veri kümesinde daha yüksek sınıflandırma doğruluğu verdiği görülmüştür.

Ilhan ve Tezel (2013), tek nükleotid polimorfizmlerinin (TNP) seçimi için parametre optimizasyonlu DVM'yi kullanmışlardır. Çalışmada, TNP'lerin tahmini DVM ile TNP taglarının seçimi ise GA tarafından yapılmıştır. DVM'nin parametreleri olan c ve γ'nin optimizasyonu için PSO kullanılmışlardır. Diğer metotlarla kıyaslandığında çalışmanın TNP tag belirlemede daha yüksek tahmin yüzdesine sahip olduğu gözlemlenmiştir.

Kuo ve ark (2013), radyo frekans tanımlama tabanlı konumlandırma için yapay bağışıklık sistemi (YBS) ve PSO ile elde edilen hibrit sistem tabanlı DVM (HBP-DVM)'yi literatüre kazandırmışlardır. Altı farklı *benchmark* fonksiyonu ile UCI makine öğrenmesi kütüphanesinden farklı veri setleri üzerinde testler yapılmış olup HBP-DVM isimli hibrit sistem, PSO-DVM ve YBS-DVM ile kıyaslanmıştır. Sonuçlara bakıldığında HBP-DVM'nin PSO-DVM ve YBS-DVM'den doğruluk oranı açısından daha yüksek başarı kaydettiği tespit edilmiştir.

Liu ve ark (2013), aqua-kültürü su kalite tahmini konusunda GA optimizasyonu ile DVM regresyonunun (GA-DVR) hibrit bir yaklaşımını önermişlerdir. Bu çalışmada gerçek değerli GA kullanılarak optimal destek vektör regresyon parametreleri araştırılmıştır. Çalışmanın sonuçlarına bakıldığında önerilen GA-DVR'nin, standart GA-DVR'den ve GBYSA'dan daha iyi performans sağladığı görülmüştür.

Peng ve ark (2013), ikili sınıflandırma problemleri için yapısal ikiz parametrik-margin destek vektör makinesini literatüre kazandırmışlardır. Önerilen bu yöntemin genelleştirme performansı açısından diğer öğrenme algoritmalarından daha başarılı olduğu görülmüştür.

Vieira ve ark (2013), septik hastalarının hayatta kalma tahmin çalışmasını yapmak için DVM'yi kullanarak ÖS'yi gerçekleştirmek amacıyla yeni bir ikili PSO (MİPSO)'yu önermişlerdir. MİPSO çeşitli *benchmark* veri setleri üzerinde test edilmiş olup, önerilen bu yöntem PSO tabanlı diğer algoritmalarla ve GA ile de karşılaştırılmıştır. Sonuçlar incelendiğinde MİPSO'nun sınıflandırma başarısının diğer yöntemlerden daha başarılı sonuç verdiği görülmüştür. Bununla birlikte alt küme çözümlerinin daha az özelliğe sahip olması da MİPSO'nun bir avantajı olarak tespit edilmiştir.

Carrizosa ve ark (2014), DVM için parametre ayarlamasında sezgisel bir nested yöntem önermişlerdir. Çalışmalarında, grid arama yönteminin uygun olmadığı *Multiple Kernel Learning* ayarlama problemleri üzerine odaklanmışlardır.

Chou ve ark (2014), sınıflandırma problemleri için hızlı bir dağınık GA kullanarak (hdGA) DVM'nin çekirdek parametrelerini optimize etmeyi hedeflemişlerdir. Bu çalışmada, hdGA parametre optimizasyonu yaparken, DVM ise eğri uydurma ve öğrenmeyi sağlamaktadır. Önerilen hdGA tabanlı DVM'nin cross-fold tahmin başarısı açısından CART, CHAID, QUEST, ve C5.0 gibi modellerden daha yüksek olduğu tespit edilmiştir.

Devos ve ark (2014), spektroskopik zeytinyağı verilerine uygulanan paralel GA'ya dayalı eş zamanlı veri ön-işlemeyi ve DVM sınıflandırma modeli seçimini literatüre kazandırmışlardır. Bu yöntem İtalyan zeytinyağının coğrafi kökeninin tespiti/ayrımı amaçlı yakın kızılötesi (NIR) ya da orta kızılötesi (FTIR) spektrumlar yardımıyla test edilmiştir.

Feng ve ark (2014), karınca koloni ağı ile DVM'yi birlikte kullanarak (KKO-DVM) izinsiz ağa girişlerin tespiti amacıyla network verileri üzerinde çalışma yapmışlardır. Önerilen yöntemi standard *benchmark KDD99* veri kümesine uygulamışlar ve sonuçları salt DVM ve KKO ile karşılaştırmışlardır. Her iki yöntemden de gerek sınıflandırma gerekse çalışma zamanı bakımından önerilen KKO-DVM'nin daha iyi olduğu görülmüştür.

Güraksın ve ark (2014), kemik yaşı belirlenmesinde PSO'ya dayanan DVM sınıflandırma yöntemini önermişlerdir. Bu işlemi gerçekleştirirken öncelikle görüntü işleme tekniğini kullanarak özellik çıkarımı yapmışlardır. Elde edilen bu özellikleri daha sonra sınıflandırıcıya giriş verisi olarak vermişlerdir. Yeni eğitim örneğini oluşturmak için de PSO Algoritmasını kullanmışlardır. Önerilen yöntem, naive Bayes, en yakın komşuluk (k-NN), DVM ve C4.5 algoritmalarıyla kıyaslandığında önerilen yöntemin diğer algoritmalardan daha iyi olduğu tespit edilmiştir.

Han ve ark (2014), GSA kullanılarak özellik alt küme seçimi (ÖAS) üzerinde çalışmışlardır. Önerilen algoritmanın, iyi bilinen benzer sınıflandırma sistemlerinden ayırt edici giriş özelliklerini seçme ve yüksek sınıflandırma doğruluğu açısından daha iyi olduğu görülmüştür.

Hsieh ve Hu (2014), DVM ile çok amaçlı PSO (ÇAPSO)'yu birlikte kullanarak geliştirilen hibrit bir sistemle parmak tanıma üzerine bir çalışma gerçekleştirmişlerdir. Önerilen sistemin uygunluğu *NIST-4* isimli veri tabanı üzerinde test edilmiştir. Sonuçların beklenildiği gibi iyi değerlere sahip olduğu gözlemlenmiştir.

Hu ve ark (2014), DVR kullanarak kısa vadede yük tahmininde model seçimi için PSO tabanlı memetik bir algoritmayı önermişlerdir. Model seçimi için kullanılan

önerilen yöntemin yararları iki gerçek dünya elektrik yük problemi üzerinde doğrulanmıştır.

Maldonado ve ark (2014), melez tamsayı programlama ile DVM için ÖS üzerine çalışma gerçekleştirmişlerdir. Önerilen yöntemde optimizasyon modeli kullanarak sınıflandırıcıyı oluşturmakla beraber eş zamanlı olarak değişken seçimini de önermişlerdir.

Miranda ve ark (2014), DVM parametre seçim problemi için ÇAPSO ile meta öğrenmeyi kombine bir şekilde kullanarak çok amaçlı hibrit bir mimari önermişlerdir. Çalışmada DVM'nin konfigürasyonuna ilişkin başlangıç Pareto yüzeyini önermek için meta öğrenmeyi kullanmışlardır. Önerilen yöntem yüz sınıflandırma problemleri olan bir veri kümesi içinde yüksek kaliteye sahip Pareto yüzeyleri elde edilerek geleneksel çok amaçlı optimizasyon algoritmalarıyla kıyaslanmıştır.

Shokri ve ark (2014), hibrit meta-sezgisel algoritmalar ile DVR'yi birlikte kullanarak üretim kalitesine ilişkin yüksek güvenilirlik tahmini üzerinde çalışmışlardır. DVR kullanılarak işlenmiş gaz yağındaki sülfürün kalitesinin tahmini için meta-sezgisel yöntemlerle önerilen yöntem karşılaştırılmıştır. Meta-sezgisel yöntemler kullanılarak optimize edilen hiper parametreler yardımıyla sülfür kalitesinin tahmini için önerilen yöntemin hem doğruluk oranında hem de hesaplama zamanında daha iyi performans sağladığı görülmüştür.

Sun ve ark (2014), DVR yaklaşımı ile belirlenen çekirdek merkezleriyle yeni bir çevrim içi adaptif çekirdek metodunu önermişlerdir. *Benchmark* veri kümeleri kullanılarak yapılan nümerik deneyler önerilen algoritmanın etkinliğini ve geçerliliğini doğrulamaktadır. Önerilen yöntemde merkezlerin hesaplanması k-ortalama küme yaklaşımı ile yapılmıştır.

Wang ve Du (2014), Tayvan'daki üreticilerin ana kart sevkiyat tahmininin gerçekleştirmek amacıyla DEA ile DVR'yi birlikte kullanmışlardır. Önerilen yöntem, Bass difüzyon ve genelleştirilmiş Bass difüzyon modeli ile karşılaştırıldığında, önerilen yöntemin model uyumu ve tahmin netliği açısından daha yüksek performans sergilediği tespit edilmiştir.

Zhai ve Jiang (2014), ağaç hedef tanıma sistemi yoluyla otomatik duyuların tespiti için yeni geliştirilmiş adaptif kaotik PSO ile eğitilmiş DVM kullanmışlardır. Önerilen PSO Algoritması DVM'nin parametrelerinin optimizasyonunu gerçekleştirmiştir. Çalışmada, hedeflere ait özelliklerin çıkartılması için temel bileşen analizi (TBA) kullanılmıştır. Ardından, özelliklere ait gereksiz ve ilgisiz bilgilerin

kaldırılması için hibrit bir ÖS metodu kullanılmıştır. Farklı gerçek ölçüm veri setleri önerilen yöntemin etkinliğini doğrulamıştır.

Zhou ve Dickerson (2014), biyo-işaretçilerin tespiti için yeni bir sınıf bağımlı ÖS metodunu kullanmışlardır. Önerilen yöntemi gerçekleştirmek için F-istatistik, maksimum ilgili İkili PSO ve sınıf bağımlı çoklu kategori sınıflandırma sistemini bir arada kullanmışlardır. Önerilen yöntemin performansını analiz etmek için sekiz gerçek kanser veri kümesi kullanılmıştır. Sonuçlar sınıf bağımlı yaklaşımların her bir kanser tipi için biyo-işaretçileri etkili bir şekilde belirleyebildiğini göstermiştir.

Chen ve ark (2015), elektrik yük tahmini yapmak için yapay zekâ ve DVM'ye dayalı yeni bir hibrit uygulama algoritması gerçekleştirmişlerdir. Deneysel mod analizi, mevsimsel düzeltme, PSO ve en küçük kareler DVM'ye dayalı yeni bir kombine tahmin modeli önermişlerdir.

García Nieto ve ark (2015), uçak motorlarının kalan yararlı ömürlerinin tahmini için hibrit bir DVM-PSO tabanlı bir model geliştirmişlerdir. Kullanılan PSO algoritması DVM eğitim aşamasında regresyon doğruluğunu önemli ölçüde etkileyen çekirdek parametresinin ayarlamasını gerçekleştirmiştir. Tahmin modelinin en temel avantajı ise makinanın önceki operasyon durumları hakkında bilgiye gerek duymamasıdır.

Harish ve ark (2015), tekrar şekil almayan dalgakıranların zarar seviye tahmini yapmak için PSO tabanlı DVM'yi literatüre kazandırmışlardır. Çalışmada, PSO-DVM'nin optimal çekirdek parametreleri PSO tarafından belirlenmiştir. Her iki model de Uygulamalı Mekanik ve Hidrolik Anabilim Dalı Deniz Yapıları Laboratuvarında yapılan deneylerden elde edilen veri setinde test edilmiştir. Sonuçlar korelasyon katsayısı, quadratik hata ve dağılım indeksi gibi istatiksel ölçümlere göre değerlendirilmiştir. Polinomal çekirdek fonksiyonuna sahip PSO-DVM'nin diğer DVM modellerinden yüksek başarı gösterdiği tespit edilmiştir.

2.2. Özellik Seçimi ile İle İlgili Literatür Çalışmaları

Chen ve ark (2011), karınca koloni optimizasyonuna (KKO) dayalı görüntü ÖS üzerinde çalışmışlardır. KKO Algoritması sınıflandırma performansı ve özellik alt küme boyutu açısından sezgisel bilgi yoluyla özellik alt kümesini seçmektedir. Çeşitli görüntüler kullanılarak elde edilen sonuçlarda daha az özellikli alt kümelerle daha iyi sınıflandırma doğruluğu elde edilmiştir.

Chuang ve ark (2011), ÖS için yayın balığı etkisi kullanılarak iyileştirilmiş ikili PSO (İPSO) üzerinde çalışmışlardır. Çalışmada, İPSO'nun performansını artırmak için yayın balığı etkisi kullanılmıştır. Bu işlem global en iyi değerin belli bir ardışık iterasyonda iyileşmeme sonucunda en kötü bireyin yerine yeni bireyin popülasyona dahil edilmesi ile gerçekleştirilmiştir. Çözümlerin kalitesinin tespiti için k-en yakın komşuluk (KNN) algoritması kullanılmıştır. Literatürden alınan 10 adet sınıflandırma problemiyle yeni önerilen İPSO karşılaştırılmıştır. Sonuçlar önerilen İPSO'nun etkili bir şekilde ÖS işlemini basitleştirdiğini göstermiştir.

Chuang ve ark (2011), çalışmalarında çadır ve lojistik isimli iki tane kaotik haritayı İPSO'ya dâhil etmişlerdir. Ortaya çıkan kaotik İPSO (KİPSO)'yu ÖS için kullanmışlardır. Önerilen sistemde sınıflandırma kesinliğinin hesabı için bir sınıflandırıcı olarak KNN'i kullanmışlardır.

Garcia ve ark (2011), gen ifade mikro dizilerinde ÖS metotları hesabı için kümelemeyle elde edilen indekslerin çok amaçlı optimizasyonu üzerinde çalışma gerçekleştirmişlerdir.

Han ve ark (2011), parametre optimizasyonuna ve ÖS'ye dayalı transformatör hata tespiti üzerinde çalışmışlardır. Çalışmada parametre optimizasyonu için GA tercih edilmiştir.

Lee ve ark (2011), parametre optimizasyonuna ve ÖS kullanarak maliyete duyarlı spam tespiti yapmak amaçlı bir yöntem geliştirmişlerdir. Saptama oranlarını maksimize etmek için, *Random Forests* kullanarak tespit modelin iki parametresini optimize etmişlerdir. Her özellik eleme evresinde parametre optimizasyonu ve genel ÖS'de sadece bir parametre optimizasyonu metotlarıyla seçilen özelliklerin optimal sayısına karar verilmiştir.

Manimala ve ark (2011), iki yeni wrapper tabanlı hibrit yumuşak hesaplama tekniğini dokuz farklı enerji arızasını sınıflandırmak amacıyla DVM'nin sınıflandırma doğruluğunu düşürmeden ÖS ve parametre optimizasyonu için önermişlerdir. Geleneksel grid arama yöntemiyle kıyasladığında, gereksiz özelliklerin veri kümesinden çıkartılması ve sınıflandırıcı için hazır parametre seçiminden dolayı GA ve ışıl işlem tabanlı yaklaşımların daha iyi sonuçlar elde etmişlerdir.

Pan ve ark (2011), İPSO ile çok ölçekli ÖS kullanarak yüz tanıma için farklı türdeki özellikleri birleştirme üzerine çalışmışlardır. ÖS yöntemi İPSO algoritmasını kullanarak her bir ölçek seviyesinde tüm arama uzayını araştırır ve en yüksek doğruluk oranına sahip minimum sayıdaki özellik alt kümeyi seçer.

Song ve ark (2011), ayrık PSO'ya ve en küçük kareler DVM (KKDVM)'ye dayalı yeni bir ÖSA önermişlerdir. Önerilen algoritma, KKDVM'yi kolay çözme ve PSO evresinde uygunluk fonksiyonlarının temel parçası olarak sınıflandırma doğruluğunu kullanma avantajlarına sahiptir. Önerilen algoritmanın sınıflandırıcıya çok büyük katkıları olabilecek özellikleri elde edebileceği simülasyon sonuçlarında görülmüştür.

Wu ve ark (2011), kaba kümeleme ve KKO algoritmasını bir arada kullanan ÖS metodunu geliştirmişlerdir. Önerilen yöntemde, özellik alt küme uzunluğu feromen güncelleme stratejisini geliştirmek için kullanılmışlardır.

Zhao ve ark (2011), DVM'nin asimptot özellikleriyle GA birleştirilerek ÖS ve parametre optimizasyonu üzerinde çalışmışlardır. Çalışmada, GA'nın arama uzayını yöneten özellik kromozomlarını üretmişlerdir. Deneysel çalışmaları *UCI*'den aldıkları gerçek dünya problemleri ve *Benchmark* veri tabanındaki veri setleri ile gerçekleştirmişlerdir. Grid arama, özellik kromozomlarına sahip olmayan GA ve diğer yaklaşımlarla kıyas edildiğinde, önerilen yöntemin yüksek sınıflandırma doğruluğuna ve daha az özellik alt kümesine sahip olduğu görülmüştür. Buna ek olarak önerilen sistemin daha az işlem süresiyle sonuca vardığı görülmüştür.

Hacıbeyoglu (2012), disjunktif normal formun indirgenmiş fark fonksiyonundan elde edilmesi yöntemini geliştirmiş ve bu yolla karmaşıklık işlemi kendi kareköküne kadar azaltılmıştır. Böylece, iki aşamadan oluşan lojik fonksiyon tabanlı bir ÖS yöntemi geliştirilmiştir. Birinci aşamada veri kümesinin doğruluk tablosu kullanılarak indirgenmiş fark fonksiyonu oluşturulur, ikinci aşamada ise elde edilen indirgenmiş fark fonksiyonu iteratif olarak bölünerek disjunktif normal forma dönüştürülür ve böylece işlenmekte olan veri kümesine ait özelliklerin en küçük alt kümeleri elde edilir. Bu özelliği sayesinde, geliştirilen ÖS yönteminin diğer ÖS yöntemleri ile işlenemeyen veri setlerini de başarıyla işleyebildiği gözlemlenmiştir.

Ahila ve ark (2012), enerji sistemleri arızalarında olasılıksal sinir ağları (OSA) için eş zamanlı model ve ÖS'yi gerçekleştiren PSO algoritmasını geliştirmişlerdir. OSA'da çok dikkat edilmeyen noktalardan birisi genişleme parametresinin seçilmesidir. PSO-OSA yaklaşımıyla yararlı özellikler seçilirken aynı zamanda OSA'nın performansını artırmak için genişleme parametresi belirlenmiştir. Deneysel çalışmalarda OSA için ayırıcı giriş özellikleriyle sınıflandırma doğruluğunun önemli ölçüde arttığı tespit edilmiştir.

Aneesh ve ark (2012), çalışmalarında parçacıkların zeki davranışlarıyla hızlandırılmış İPSO'yu önermişlerdir. Çözünürlük dönüşümü, histogram eşitleme ve kenar bulma gibi görüntü ön işleme teknikleriyle birlikte hızlandırılmış İPSO yöntemini tanıma oranını artırmak ve ÖS için özellik alt kümesini azaltmak amacıyla kullanılmıştır.

Cervante ve ark (2012), İPSO'ya ve bilgi teorisine dayanan sınıflandırma problemleri için iki yeni filtreleme ÖS yöntemi geliştirmişlerdir. İlk algoritma PSO'ya ve seçilen özellik alt kümesinin bağlantısını belirleyen her bir çift özelliğin karşılıklı bilgisine dayanır. İkinci algoritma ise PSO'ya ve her bir grup özelliğin entropisine dayanır. Sonuçlar incelendiğinde, ilk algoritmanın daha az özelliğe sahip alt küme bulurken ikincisinin de daha yüksek başarı kaydettiği görülmüştür.

Adam ve ark (2014) , EEG sinyallerinde zirve tespiti amacıyla PSO algoritması kullanarak ÖS ve sınıflandırıcı parametre tahmini üzerinde çalışmışlardır. Çalışmada standart PSO ve random asenkron PSO olmak üzere iki farklı PSO modeli kullanılmıştır. Random asenkron PSO'nun standart PSO'ya göre daha iyi ve güvenli sınıflandırma oranı verdiği gözlemlenmiştir.

De la Hoz ve ark (2014), çok amaçlı optimizasyon ile ÖS üzerinde yaptıkları çalışmada kendini organize eden hiyerarşik haritalar yardımıyla ağ içindeki kusurların tespiti amaçlı bir uygulama geliştirmişlerdir. Önerilen yöntemin performans analizini görmek için *DARPA/NSL-KDD* veri kümesi kullanılmıştır.

Erguzel ve ark (2014), teta ve delta frekans bantlarından 6 kanal tedavi öncesi verileri için GBYSA sınıflandırma metodu ile KKO algoritmasını hibrit kullanmışlardır.

Forsati ve ark (2014), çalışmalarında KKO algoritmasının yeni bir versiyonu olan zenginleştirilmiş KKO (ZKKO) algoritmasını geliştirmişler ve uygulamasını da ÖS üzerinde yapmışlardır. Önerilen algoritmanın etkinliğini göstermek için çeşitli standart veri setleri üzerinde kappa testi ve sınıflandırma doğruluğu şeklinde iki farklı ölçüm kullanarak deneysel çalışmalar yapılmıştır.

Ghamisi ve Benediktsson (2014), Salinas hiperspektral veri kümesi üzerinde yaptıkları çalışmada GA ve PSO'nun birlikte kullanımıyla ÖS'yi gerçekleştirmişlerdir. Önerilen yöntemin uygun bir CPU zamanı içerisinde en öğretici özellikleri otomatik olarak seçebildiği doğrulanmıştır.

Han ve ark (2014), modifiye edilmiş GSA (MGSA) ile ÖAS yöntemini geliştirmişlerdir. MGSA'nın başarısındaki temel nokta türlerin çeşitliliğini artırmak için parçalı lineer kaotik bir harita kullanılması ve lokal kullanımı artırmak için sıralı

quadratik programlamanın kullanılmış olmasıdır. MGSA CEC 2015 özel oturumdan elde edilen 10 fonksiyon üzerinde test edilmiştir. Sonuçlar farklı MGSA, PSO ve GA karşılaştırılmıştır. Elde edilen sonuçlar önerilen MGSA'nın optimizasyon alanında çeşitli problemleri çözmede daha yüksek performans sergilediğini göstermiştir.

Olfati ve ark (2014), Wisconsin göğüs kanseri teşhisinde DVM için ÖAS ve parametre optimizasyonu üzerinde yaptıkları çalışmada özellik azaltma için TBA, ÖS için GA ve sınıflandırma için de DVM'yi kullanmışlardır. Bu yaklaşım özellik sayısını minimize etmenin yanında duyarlılık, spesifiklik ve alıcı işlem karakteristik eğrisini maksimuma taşıyacak özelliklere sahip optimal sınıflandırmayı sağlamıştır. Test aşamasında 10 fold çapraz geçerleme yapılmıştır. Geliştirilen TBA+GA+DVM sistemi iki özellik içeren alt küme için %100 başarı elde etmiştir.

Saha ve ark (2014), ısıl işlem tabanlı çok amaçlı optimizasyon (IİÇAO) kullanarak yarı denetimli öğrenme ve ÖS ile ilgili problemleri çözmüşlerdir. Önerilen IİÇAO yöntemi verilen herhangi bir veri setinde uygun bölünmenin yanı sıra uygun sayıda kümeleme ve uygun özellik alt kümelerini belirlemek için kullanılmıştır. Sarac ve Ozel (2014), web sayfalarının sınıflandırma doğruluğunu artırmak ve çalışma süresini iyileştirmek amacıyla özellik sayısını azaltacak KKO algoritması tabanlı bir ÖS yöntemi geliştirmişlerdir. Deneylerde *WebKB* ve *Conference* veri setleri tercih edilmiştir. KKO tabanlı ÖS'nin bilgi kazancı ve chi kare ÖS yöntemlerine göre daha iyi sonuçlar verdiği görülmüştür.

Wang ve ark (2014), KİPSO algoritması, korelasyon analizi ve özellik derecelendirme kullanarak sınıflandırma performansının iyileştirilmesi ve gereksiz özelliklerin azaltılmasına odaklanan çok aşamalı bir ÖSA geliştirmişlerdir. İlk aşamada, özellik derecelendirme ile sınıflandırma doğruluğuna göre derecesi en yüksek olan özellikler seçilmiştir. İkinci evrede, gereksiz özelliklerin çıkartılması amacıyla en yüksek başarı derecesine sahip seçilen özellikler arasında korelasyonu ölçmek için korelasyon analizi kullanılmıştır. Son evrede ise KİPSO ile optimal alt küme seçimi yapılmıştır. N-kat çapraz doğrulama yapılarak DVM 6 farklı veri kümesi üzerinde deneysel çalışmaya tabi tutulmuştur.

Wang ve ark (2014), boyut azaltmak amacıyla bakteri kolonisi algoritmasına (BKA) dayalı yeni bir ağırlıklandırılmış ÖS geliştirmişlerdir. Ağırlıklandırma stratejisi aynı bireyde tekrar eden görünümün yanı sıra BKO ile seçilen özelliklerin frekansını dikkate alan bir stratejidir. Bununla birlikte öğrenme stratejisini oluştururken görünmeyen özelliklere ait bilgisizlik tarafını engelleyen rastsallığı dikkate almışlardır.

Önerilen yöntemin etkinliğini görmek için farklı boyutlardaki *benchmark* fonksiyonları ile testler yapılmıştır. Önerilen ağırlıklandırılmış ÖSA son zamanlarda önerilmiş olan üç farklı popülasyon tabanlı ÖSA ile karşılaştırılmış ve önerilen sistemin başarısı doğrulanmıştır.

Xue ve ark (2014), arama aşamasında elde edilen daha önemli çözümleri depolayacak harici bir arşive sahip yeni bir PSO tabanlı ÖSA'yı geliştirmişlerdir. Önerilen yöntemin iki özel metodu bulunmaktadır. PSOArR arşivden *gbest* değerini rastsal olarak seçmektedir. PSOArRWS ise hem özellik sayısını hem de sınıflandırma hata oranını düşünerek *gbest* değerini seçmek için rulet tekerleği seçim yöntemini kullanmaktadır. 12 farklı *benchmark* fonksiyonu üzerinde yapılan deneysel çalışmalarda PSOArR ve PSOArRWS'nin tüm özellikler kullanılarak elde edilen başarıdan daha yüksek başarı elde ettikleri görülmüştür.

Emary ve ark (2015), optimal özellik alt kümesini bulmak için gri kurt optimizasyon yöntemiyle sınıflandırma doğruluk tabanlı uygunluk fonksiyonunu geliştirmişlerdir. Bir dizi UCI makine öğrenmesi veri kümesi üzerinde yapılan deneylerde, GA ve PSO ile kıyaslandığında hem sınıflandırma doğruluğunda hem de özellik boyut azaltmada önerilen yöntemin daha iyi sonuçlar verdiği tespit edilmiştir. Bunun dışında algoritmanın çok kararlı olduğu yapılan testlerde görülmüştür.

Banka ve Dara (2015), yüksek boyutlu ÖS, sınıflandırma ve validasyon yapmak amacıyla *hamming* uzaklık tabanlı İPSO algoritmasını geliştirmişlerdir. Önerilen algoritmanın verimliliğini ve üstünlüğünü göstermek için üç farklı *benchmark* veri kümesi üzerinde deneysel çalışmalar detaylıca yapılmıştır. Çalışma sonucunda, hazır ön işleme metotları seçerek önerilen İPSO yönteminin gen dizilimlerinde önemli özellikleri bulmayı mümkün kılacağı görülmüştür.

Ghamisi ve Benediktsson (2015), PSO ve GA'nın birleştirilmesi ile oluşan hibrit sisteme dayalı yeni bir ÖS yaklaşımı geliştirmişlerdir. Deneylerini Indian Pines Hiperspektral veri kümesi üzerinde gerçekleştirmişlerdir. Ayrıca önerilen yöntem yol tayini için de test edilmiş ve sonuçlar arka plan ve yol arasındaki farkı bulmada önerilen hibrit sistemin yetenekli olduğu tespit edilmiştir.

Lin ve ark (2015), yapay balık koloni algoritması (YBKA)'ın lokal minimuma takılma ve çeşitlilik eksikliği gibi dezavantajlarından dolayı çalışmalarında modifiye edilmiş yapay balık koloni algoritmasını (MYBKA) kullanmışlardır. MYBKA'ya dayalı DVM için ÖS ve parametre optimizasyonu üzerinde çalışmışlardır. Bilinen UCI veri

setleri üzerinde yapılan deneysel sonuçlarda daha az özellikli alt kümeler kullanarak sınıflandırma doğruluğu bakımından MYBKA'nın üstünlüğü görülmüştür.

Lin ve ark (2015), standart kedi sürüsü optimizasyonu algoritmasını (KSOA) DVM'nin parametre optimizasyonunu ve ÖS'yi lokal bir arama prosedürü ile bütünleştirmişlerdir. MKSOA ile KSOA kıyaslandığında daha az özellikle oluşturulan alt küme kullanılarak elde edilen sınıflandırma doğruluğunda MKSOA'nın daha üstün olduğu gözlemlenmiştir.

Liu ve ark (2015), moleküler imzada ÖS için ayrık biyojeografi tabanlı optimizasyon (ABTO) yöntemini geliştirmişlerdir. Çalışmada ilk olarak, Fisher-Markov seçim modelini sabit sayıdaki gen verilerini seçmek için kullanmışlardır. İkinci olarak ise, önerilen yöntemi ÖS için uygun formata getirmede algoritmanın araştırma ve kullanma yeteneğini dengelemek için ayrık mutasyon ve ayrık göç modelini kullanmışlardır. UCI'den alınan dört göğüs kanser veri kümesi ile deneysel çalışmaları yapmışlardır. Önerilen ABTO yöntemi, GA, PSO ve DEA ile kıyaslandığında ABTO'nun literatürdeki önceki yöntemlere göre daha iyi sonuç verdiği görülmüştür.

Liu ve ark (2015), HIV-1 proteazının özgünlüğünü analiz etmek ve oktapeptitlerdeki konumları bulmak amacıyla YSA optimizasyonu ile kombine edilmiş yeni bir ÖS yöntemi geliştirmişlerdir.

Moayedikia ve ark (2015), her bir yemek kaynağı için lokal ve global ağırlıkları dikkate alan ağırlıklandırılmış arı koloni optimizasyon (AAKO) algoritmasını geliştirerek ayrık problemlere uygulamışlardır. Ayrıca AAKO'yu kullanarak ÖS için bir uygulama geliştirmişlerdir. AAKO algoritmasının performansını göstermek için standart *benchmark* fonksiyonları kullanılarak yapılan deneylerde önerilen AAKO yönteminin yüksek performans sergilediği gözlemlenmiştir.

Moradi ve Rostami (2015), sınıflandırma problemlerini çözmek için üç aşamadan oluşan graf kümeleme yaklaşımına ve KKO algoritmasına dayalı yeni bir ÖS yöntemi geliştirmişlerdir. Bu yaklaşımların ilkinde, tüm özellik kümesini bir graf olarak temsil etmişlerdir. İkinci aşamada, bir ağ belirleme algoritması kullanılarak özellikleri belli bazı gruplara bölmüşler ve son olarak da üçüncü evrede, nihai özellik alt kümesini seçmek için KKO algoritmasına dayalı yeni bir arama stratejisi geliştirmişlerdir.

2.3. Ayrıklaştırma ile İle İlgili Literatür Çalışmaları

Jiang ve ark (2009), tahmini EFD modeli üzerinde çalışmışlardır. Önerilen bu modeli UCI veri tabanındaki veri setleri üzerinde diğer ayrıklaştırma yöntemleri ile kıyasladıklarında önerilen yöntemin etkili ve kullanışlı olduğunu görmüşlerdir.

Orhan ve ark (2012), EEG verilerinden epilepsi krizinin tespiti amacıyla EFD'ye dayalı olasılık dağılım adı verilen yeni bir özellik çıkarım yaklaşımını önermişlerdir. İki farklı olasılıksal yoğunluk fonksiyonu için ortalama karesel hata kriteri kullanıldığında epilepsi kriz tespiti üzerinde % 96.72 oranında bir başarı sergilemişlerdir. Buna ek olarak, EEG segmentinin olasılıksal yoğunluğu çok katmanlı YSA modelinin girişi olarak kullanıldığında epilepsi krizi saptama başarısının % 99.23'e çıktığı gözlemlenmiştir.

Koçoğlu (2012), veri kümelerine farklı ayrıklaştırma yöntemlerinin uygulanmasını ve hangi yöntemin daha etkin olduğu konusunun incelenmesini amaçlamıştır. Çalışma kapsamında Wisconsin veri kümesi kullanılmıştır. Bu veri kümesi üzerine KEEL veri madenciliği yazılım aracı yardımı ile 1RD, CADD, CAIM, Chi2, ChiMerge, ID3, Eşit Genişlikli, Eşit Frekanslı olmak üzere toplamda sekiz farklı ayrıklaştırma yöntemi uygulamıştır. Deneysel sonuçlar incelendiğinde; Chi2, ChiMerge, CAIM algoritmalarının önemli oranda tutarlı çalıştıkları, 1RD algoritmasının genelde bir, ID3 algoritmasının da çok sayıda kategorik değişken elde ettiği gözlemlenmiştir.

Doğan (2012), doktora tez çalışmasında ayrıklaştırma yöntemleri ve YSA kullanarak asenkron motorlarda arıza teşhisi üzerinde çalışmıştır. Çalışmasında, asenkron motorların rulman, eksenden kaçıklık, rotor çubuk kırığı ve stator sargı kısa devresi arızalarının tespitine odaklanmıştır. Bir asenkron motorun arızalarına ilişkin bu karakteristik özellikleri, motora ait akım sinyallerinin analizini kullanarak çıkarmıştır. Sinyal analiz yöntemi olarak önceki çalışmalarda yaygın olarak kullanılan frekans ve zaman-frekans yöntemleri yerine, zaman boyutunda ayrıklaştırmaya dayalı bir yöntem kullanmıştır. Gerçekleştirilen deneysel çalışmalarda farklı yük koşullarında sağlam ve arızalı durumlar için motor akım sinyalleri ölçülmüştür. Bu akım verisi zaman boyutunda EWD ve EFD yöntemleri ile ayrıklaştırılarak özellikler çıkarılmıştır. Bu özellikler ile eğitilen çok katmanlı algılayıcı YSA kullanılarak bu arızaların tespiti ve sınıflandırması gerçekleştirilmiştir. Sonuç olarak önerilen yöntemin asenkron motor arızalarının tespitinde oldukça başarılı sonuçlar verdiği görülmüştür.

Chaves ve ark (2013), emisyon bilgisayarlı tomografi görüntülerine dayalı Alzheimer hastalarının erken teşhisi için sürekli özellikleri ayrıklaştırma ve ilişkili kural çıkarma yöntemlerini bir arada kullanarak bilgisayar destekli bir teşhis sistemi geliştirmişlerdir.

Sayın (2013), veri madenciliğinin etkisini kullanarak, ÖS ve ayrıklaştırma yöntemleriyle beraber sınıflandırma algoritmalarının kullanılmasıyla koroner arterlerde kalsifikasyon olup olmadığının incelenmesinin önemini göstermiştir. Veri setinde yer alan yaş, nakil süresi, diabetik ve fosfor gibi özelliklere bakarak doktorların bir hastada kalsifikasyon olup olmadığına yaklaşık olarak %70 doğrulukta karar verebileceği bir sistem geliştirmiştir. Veri ayrıklaştırma işleminden sonra ise bu oranın %75 civarına yükseldiği gözlemlenmiştir.

Ozdemir (2014), tespit sistemlerinde kullanılan bir çok makine öğrenmesi tekniğini detaylıca incelemiştir. Başarılı bir hizmet engelleme saldırısı tespit oranına ulaşmak amacıyla, yerine geri koyarak örnekleme, ayrıklaştırma, öznitelik seçimi ve sınıflandırma yöntemlerinin birleşimini temel alan bir saldırı tespit sistemi önermiştir. Tez çalışmasında ayrıca hizmet engelleme saldırılarının tespiti üzerinde çalışmakla beraber çeşitli ağ saldırılarını ve normal ağ trafiğini doğru olarak tespit etmekle de çalışmıştır. Ayrıklaştırma yöntemi olarak; düzensiz temelli ayrıklaştırma, eşit genişlikli gruplama, eşit frekanslı gruplama ve orantısal k aralıklı yöntemlerini kullanmıştır. Karmaşıklığı azaltmak için ise, korelasyon temelli ÖS, tutarlılık temelli ÖS, bilgi kazancı temelli ÖS ve simetrik belirsizlik temelli ÖS yöntemlerini kullanmıştır. Deneysel çalışmalar için de KDD90 eğitme ve test etme verileri kullanılmıştır.

Jedrzejowicz ve Skakovski (2014), ayrık-sürekli değerlere sahip çizelgeleme problemlerini çözmek için ada tabanlı DEA'yı geliştirmişlerdir. Önerilen algoritmanın orijinal DEA'dan daha iyi çözümler bulduğu tespit edilmiştir.

Bruni ve Bianchi (2015), ayrıklaştırma ve istatistiksel analizlere dayalı küçük bir eğitim seti kullanarak etkili bir sınıflandırma yöntemi geliştirmişlerdir. Önerilen yaklaşım veri kümesi üzerinde istatistiksel faktörlerden elde edilen bilgilerle zenginleştirilerek mantıksal analiz verisi (LAD) çatısı üzerine inşa edilmiştir. Önerilen yöntem UCI veri kümesinden alınan veri setleri üzerinde standart LAD, DVM ve etiket yayılım algoritmaları ile kıyaslanmış ve sonuçların gözle görülür oranda iyi olduğu tespitine varılmıştır.

Akar ve ark (2015), çok katmanlı algılayıcı YSA modeli ve EWD'ye dayalı olasılıksal dağılımları kullanarak kalıcı mıknatıs senkron motorlarının kusurlarını

çözmeye ve statik dışmerkezliliği tespit etmeye odanlanmışlardır. Veriler EGA sayesinde birkaç belli aralığa bölünmüştür. Olasılıksal dağılım elde edilen bu ayrıklaştırılmış değerlere ait her bir aralıktaki veri sayısına göre hesaplanmıştır ve bu dağılımlar çok katmanlı YSA modeline giriş verisi olarak kullanılmıştır. Önerilen yöntem kusur saptama performansının ölçülmesi için 18 farklı şekilde deney yapılmıştır. Önerilen yöntemin bazı deneylerde yüksek hız ve tam yüklemede çok başarılı olduğu gözlemlenmiştir.

3. OPTİMİZASYON ALGORİTMALARI

Optimizasyon, bir problemde belirli kısıtlar altında mevcut olan alternatif çözümler içinden en iyisini bulma işlemi olarak tanımlanmaktadır. Literatürde optimizasyon problemleri için birçok algoritma önerilmiştir. Sezgisel algoritmalar, büyük boyutlu optimizasyon problemleri için, belli bir süre içerisinde en uygun çözüme yakın sonuçlar üretebilen algoritmalardır. Genel amaçlı sezgisel optimizasyon algoritmaları, biyoloji tabanlı, fizik tabanlı, sürü tabanlı, sosyal tabanlı, müzik tabanlı ve kimya tabanlı olmak üzere altı farklı gruba ayrılmaktadır (Akyol ve Alataş 2012). Bu tez çalışmasında sürü tabanlı algoritma grubundan olan ABC, BA, CSO algoritmaları kullanılırken, fizik tabanlı algoritma grubundan ise GSA tercih edilmiştir.

3.1. Yapay Arı Koloni Algoritması

Doğada bulunan bal arıların besin ararken kullandıkları yöntemlerden esinlenerek geliştirilmiş olan ABC algoritması, sürü zekâsına dayalı bir optimizasyon algoritmasıdır (Karaboga 2005). Doğada sürü şeklinde hareket eden arıların besin bulmada ortaya koymuş oldukları davranışlardan esinlenilen bu algoritma sayesinde genellikle optimizasyon problemlerini çözmek hedeflenmektedir.

3.1.1. Arıların yem bulma davranışları

Arı kolonisinin yaşamının devamı için en önemli işlemden birisi besin arama işlemidir. Bu süreçteki önemli parametreler ise; arılar tarafından kovanda biriktirilen kaynaklar, ortamdaki yemek kaynakları ve arıların birbirleriyle kurdukları haberleşme ağı olarak bilinir.

Arama işlemi arının kovandan ayrılmasıyla başlar ve bu süreci rastgele yapılan yiyecek aramaları takip eder. Yiyecek aramaya çıkmış olan arılar, buldukları kaynaklarda yiyecek miktarının azaldığını fark ettiklerinde, diğer arılardan aldığı bilgiye göre başka kaynaklara hareket ederler. Bu aşamada bulunan kaynaklara ait bilgilerin arılar arasında iletilmesi ve ortamdaki polen, su gibi kaynakların kovana getirilmesi işlemleri yapılır.

Tereshko'nun öngördüğü yiyecek arama modelinde üç temel öğe bulunmaktadır. Bunlar: yiyecek kaynakları, görevli arayıcılar, görevi belli olmayan arayıcılardır (Tereshko ve Loengarov 2005).

Akay, çalışmasında yiyecek kaynakları, görevli arayıcılar, görevi belli olmayan arayıcıları aşağıdaki gibi açıklamıştır (Akay 2009):

3.1.3. Yiyecek kaynakları

Arıların yiyecek bulmak için gittiği kaynaklar yiyecek kaynakları olarak isimlendirilir. Doğada bulunan bir yiyecek kaynağının değerini, kaynağın çeşidi, kaynağın yuvaya olan uzaklığı ve kaynaktaki nektar miktarı gibi birçok etken belirlemesine rağmen daha basit modellenmesi için sadece kaynağın zenginliği parametre olarak kabul edilebilir.

3.1.3. Görevi belirli işçi arılar

Daha önceden belirlenmiş kaynaklardan toplanan yiyeceğin kovana getirilmesi işini işçi arılar yerine getirmektedir. İşçi arıların bir diğer önemli görevi de gittikleri kaynağın konumu ve kaynağın kalitesi hakkında diğer arılara bilgilendirme yapmaktır.

3.1.4. Görevi belirli olmayan işçi arılar

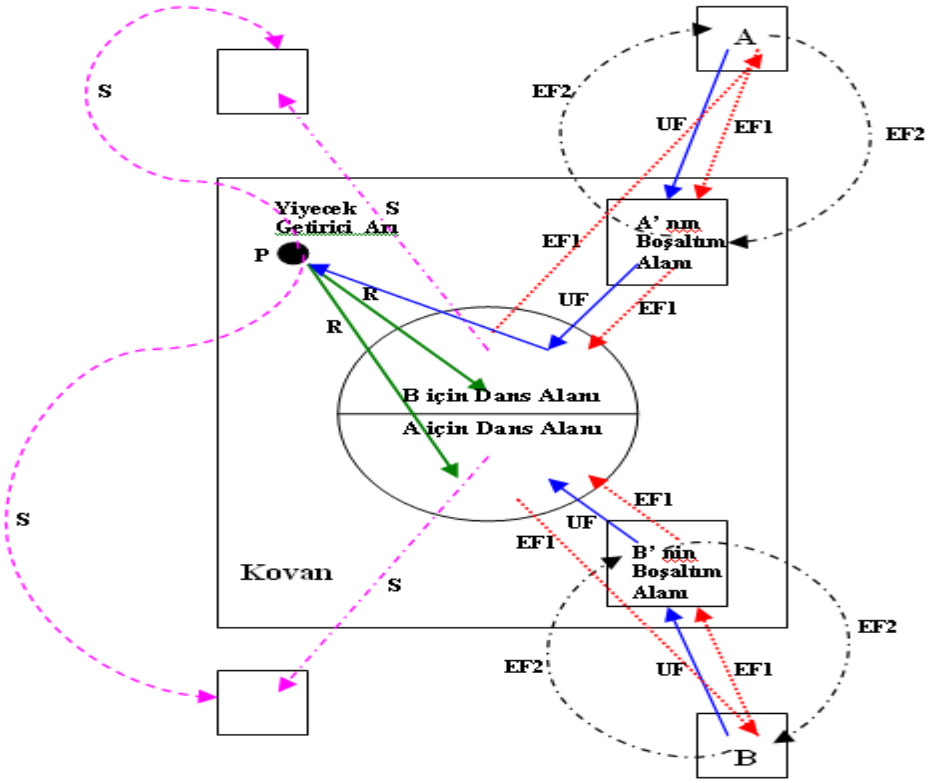
Bu arıların yiyecek olma ihtimâli olan kaynakları arama görevleri vardır. Görevi belli olmayan iki tip işçi arı vardır. Bunlar; kâşif ve gözcü arılardır. Kâşif arılar arama bölgesinde rastgele kaynak arayan arılardır. Gözcü arılar ise kovanda bekleyen ve görevli arıları izleyen arılardır. Bunun sonucunda görevli arılar tarafından gelen bilgiyi dikkate alarak yeni kaynaklara hareket ederler.

Birliktelik yapısının ve ortak bilginin oluşmasını sağlayan en önemli konu arıların birbirleri arasındaki bilgi paylaşımı konusudur. Arılar için yaşam alanı olarak tanımlanan kovan bazı bölümlere ayrılabilir. Bilgi paylaşımının yapıldığı dans alanı bu bölümlerden bir tanesidir. Arılar yapmış olduğu dansla kendi aralarında bilgi paylaşımını sağlarlar. Paylaşılan bu bilgi sayesinde kaliteli yeni besin kaynakları bulunmuş olur (Gruter ve Farina 2009).

Kaynaklardan yiyecek getirmiş olan arıların, diğer arıların bu kaynaklara yönlendirilmesi için kaynağın konum bilgisini diğer arılara vermesi gerekir. Konum hakkında bilgiyi alan arı güneş ışığını kullanarak hedefine ulaşmış olur. Arılar sahip oldukları özellikleri gereği kendi yörüngeleri ile güneş arasındaki açıyı hesaplayabilmektedirler. Uzaklıklarını tükettikleri enerjiye göre ayarlayan arılar,

taşıdıkları yüklere göre farklı yükseklikte uçarak enerjilerini ayarlamaktadırlar. (Akay 2009).

Arıların besin arama davranışları Şekil 3.1'de gösterilmiştir. Bu davranışların şekil üzerinde incelenmesini ve açıklamalarını Akay şu şekilde gerçekleştirmiştir: Şekildeki *A ve B bulunmuş kaynaklar* olarak kabul edilmiştir. Başlangıçta görevi belli olmayan ve yiyecek kaynağı bilgisine sahip olmayan arı yiyecek aramaya başlayacaktır. Bu arı için iki durumdan bahsedilebilir: Birincisi; Şekil 3.1'de S ile gösterilen bu arı kâşif arı olduğu varsayılırsa bu durumda bu arı besin aramaya başlayabilir. İkincisi ise şekilde R ile gösterilen ve dans eden arıları izleyerek yönlendirilen kaynaklara giden bir gözcü arı olabilir.



Şekil 3.1. Arıların yem arama davranışları (Akay 2009)

Yiyecek için kaynaklara giden arılar kaynaklardan nektar getirmeye başlarlar. Bu durumda bu arılar görevli arıya dönüşmüş olurlar. Bu arılar için bundan sonra üç seçenek vardır:

i) Diğer arılarla herhangi bir bilgi paylaşımı yapmadan kaynaktan nektar getirme işlemine devam edebilir. Bu arı şekilde EF2 ile gösterilmiştir.

ii) Bu arılar kaynağa varmadan dans işlemi yaparak nektar kaynağının konumu ve nektar miktarı hakkında diğer arılara bilgi paylaşımı yaparak arıları bu kaynağa yönlendirme yapabilir. Bu arı EF1 ile gösterilmiştir.

iii) Kaynağı terk edip dans alanında gözcü arı olabilir. Bu arı UF ile gösterilmiştir.

3.1.5. Yapay arı kolonisi algoritması ve temel adımları

Yapay Arı Kolonisi algoritması doğadaki arı kolonilerinin zeki davranışları ve besin arama aşamasındaki davranışları modellenerek Karaboğa tarafından geliştirilmiştir (Karaboga 2005). Algoritmada görevli arıların sayısı ile toplam yiyecek kaynağı birbirine eşittir. İşçi arıların sayısı da gözcü arıların sayısına eşittir. Kaynaktaki nektar miktarı bittiğinde görevli arı kâşif arıya dönüşmektedir. Kolonideki yemek kaynaklarının konumları problemin muhtemel çözümlerini ifade eder. Nektar miktarı ise çözümün kalite ölçüsünü ifade etmektedir. ABC algoritmasının hedefi, en çok nektara sahip yiyecek kaynağının konumunu bulmayı hedefler. Bunun sonucu olarak arama uzayındaki çözümlerden problemin en düşük ya da en yüksek değerini veren çözümü bulmaya çalışır (Akay 2009).

Yiyecek arama işleminde kâşif arılar yiyecek aramaya ilk etapta rastgele başlarlar. Kaynağı bulan kâşif arı görevli arıya dönüşür ve bu görevli arılar artık kovana nektar götürmeye başlarlar. Görevli arı kovana nektar taşır ve kendine özgü dansı ile yiyecek bulduğu kaynakla ilgili bilgiyi bekleyen gözcü arıya iletir. Kaynaktaki nektar bittiği anda görevli arılar da gözcü arıya dönüşür. Gözcü arılar dansları izledikten sonra yiyeceklerin kalite ölçüsüne göre bir kaynağa gider. Çizelge 3.1'de ABC algoritmasının adımları verilmiştir.

Çizelge 3.1. ABC algoritmasının adımları

- Yiyecek kaynağı bölgelerinin (noktalarının) oluşturulması
- TEKRARLA
    - İşçi arıları yiyecek kaynağına gönder ve nektar miktarının hesapla
    - Gözcü arıların seçim işleminde kullanacakları olasılık değerlerini hesapla
    - Gözcü arıların hesaplanan olasılık değerleri doğrultusunda yiyecek kaynağı bölgesi seç
    - Kaynaktan ayrılma kriteri için limit ve kâşif arı üret
- Çevrim sayısı maksimum çevrim sayısından küçük olduğu sürece

3.1.6. İkili (Binary) yapay arı koloni algoritması

Yapay arı koloni algoritması çoklu değişken ve çok türlü sürekli fonksiyonları çözmek için kolayca uygulanabilen sürü zekasına dayalı bir optimizasyon algoritmasıdır. ABC Algoritma performansı incelendiğinde, daha az kontrol parametresi kullanma ve hızlı yakınsama özelliği olan diğer sürü zekâsına sahip algoritmalara göre ABC'nin daha başarılı olduğu görülmüştür. Algoritmanın basitlik ve kolayca uyarlanabilirlik gibi avantajlarından dolayı ABC çoğu optimizasyon problemlerini çözmek için kullanılmıştır. ABC kısıtlı optimizasyon problemleri, sinir hücreleri ve kümeleme gibi bazı optimizasyon problemlerine başarılı bir şekilde uygulanmaktadır (Karaboga ve Basturk 2007).

Yukarıda bahsedildiği gibi, ABC algoritmasının orijinal versiyonu sürekli değerleri optimize etmek için geliştirilmiştir. Fakat gerçek dünyada çoğu optimizasyon problemi değişken düzeyi arasındaki ve değişkenler arasındaki niteliksel ve ayrık farklılıklara sahip bir uzay kümesi içinde tanımlanır. Klasik örnekleri ise çizelgeleme ve routing problemleri gibi ayrık elemanların düzenlenmesi ve sıralanmasını gerektiren problemlerdir. Araştırmacılar, karmaşık olmayan bu kombinasyonlu problemlerin yanında sıklıkla problemleri daha etkin çözüme kavuşturmak için problemleri ikili ifadelerde ondalık tipli şekle dönüştürürler ve bunları kesikli sayı uzayında çözmeye çalışırlar.

Yeni ikili ABC algoritması, parametreleri ayarlama, işçi ve gözcü arılar sayesinde elde edilen pozisyon güncellemesi ve uygunluk fonksiyonunda herhangi bir iyileşme olmadığında kaşif arıların sisteme dahil edilmesi aşamalarından oluşur.

3.1.6.1. Başlangıç evresi

Önerilen modelde, arılar tarafından kullanılan ikili notasyonda temsil edilen yiyecek kaynakları ilgili optimizasyon problemine ilişkin muhtemel çözümleri ifade eder. Her bir arıya ait olan nektar miktarı ise uygunluk fonksiyonunu temsil eder. Başlangıçta, bir dizi yemek kaynağının konumları arılar tarafından rastgele belirlenir. Yani, arı kolonisi denklem 3.1'deki gibi arama uzayında rastsal şekilde başlatılır.

$$X_{ij}(t=0) = LB_j + (UB_j - LB_j) \times rand[0,1] \qquad (3.1)$$
$$for\ i = 1,2,\dots,N\ ve\ j = 1,2,\dots,D$$

Buradaki N sayısı kolonideki arı sayısını; D arama uzayının boyutunu; LB_j ve UB_j ikili uzayda sırasıyla 0 ve 1 olarak atanan j boyutu için alt ve üst sınırları ifade etmektedir.. *rand[0,1]* ise 0 ve 1 arasında rastgele olarak seçilen ondalık bir değerdir. *i.* arı tarafından keşfedilen x_i yemek kaynağının konumundaki her bit tamamıyla rastgele seçilir (Wei ve Hanning 2012).

3.1.6.2 İşçi arı evresi

Tüm arılar kovana bir miktar nektarla geldikten sonra, en iyi yiyecek kaynağını bulan koloninin ilk yarısı *işçi arı*ya dönüşür. Bu evrede, x_i ismindeki her bir işçi arı için denklem 3.2'de görüldüğü gibi şu anki konum komşuluğunda v_i isminde yeni bir yemek kaynağı üretilir.

$$V_{ij}(t) = x_{ij}(t-1) + \varphi\left(x_{kj}(t-1) - x_{ij}(t-1)\right) \quad (3.2)$$
$$\text{for j} = 1,2,\dots,\text{D}$$

Buradaki $k \in (1,2\dots,N)$ ve $j \in (1,2,\dots,D)$ random olarak seçilen indekslerdir ve k değeri i değerine eşit olmamalıdır. φ_{ij} (-1,1) aralığında üretilen random bir sayıdır.

Arının pozisyonunu [0,1] aralığında sınırlandırmak için, normalizasyon metodunun kullanılması gerekmektedir. Ayrıca v_{ij}'nin tüm gerçek değerli sayılarını [0,1] aralığına dönüştürmek için eşik seviyesinin başlatılması gerekir. Pozisyon kırpma sınır durumu ve yuvarlama fonksiyonu bu son modifikasyonu gerçekleştirmek için kullanılabilir. Pozisyon güncelleme işlemi denklem 3.3'deki kurala göre yapılır.

$$x_{ij}^{e}(t) = round(v_{ij}(t))$$
$$\text{if } x_{ij}^{e}(\text{t}) \geq UB, ise\ x_{ij}^{e}(t) = UB \quad (3.3)$$
$$\text{else if } x_{ij}^{e}(\text{t}) \leq \text{LB}, \text{ise } x_{ij}^{e}(t) = LB$$

Buradaki round fonksiyonu v_{ij} değerini en yakın tamsayıya yuvarlar ve *if else* ifadeleri her bir boyuttaki elamanları probleme ait uzay aralığına sınırlandırır. $X_i(t)$ isimli yeni bir yemek kaynağı elde edildikten sonra uygunluk fonksiyonu hesaplanır ve $x_i(t-1)$ ile karşılaştırılır. $x_i(t)$'nin uygunluk değeri $x_i(t-1)$'inkine eşit ya da ondan daha iyi ise; $x_i(t)$ ile $x_i(t-1)$ yer değiştirir. Son durumda $x_i(t)$ popülasyonun yeni üyesi olur. Aksi

halde $x_i(t)$ o şekilde kalır. Kısaca; önceki ve şu anki yemek kaynakları arasındaki seçim operasyonu aç gözlü seçim mekanizması ile yapılır. İABC algoritmasının prosedürel kodu ise Çizelge 3.2'de gösterilmektedir (Wei ve Hanning 2012).

Çizelge 3.2. İABC algoritmasının prosedürel kodu (Wei ve Hanning 2012)

```
procedure İABC
2*NS kadar random ikili çözüm üret  X1, X2,…., X2NS
best_sol := X1, X2,…., XNS  arasında en iyi çözümü bul
tekrarla
        for i:=1 to NS do
                for j:=1 to D do
                        /* İşçi Arı Evresi */
                        Xe:= Komsu_Cozumleri_Uret(Xij)
                        if (Xe, Xij daha kötü) ise
                        Xij yi değiştirme
                        else if (Xe, Xij daha iyi) ise
                        Xij := Xe
                end for
        if (Xi Limit iterasyonuna kadar değişmedi) ise
        /* Kaşif Arı */
        Xi random çözümle yer değiştir.
        if (Xi, best_sol dan daha iyi) ise
        best_sol := Xi
        end for

        for i:= 1 to NS do
        ki := Index_Sec(X1, X2,…., XNS)
        /* Gözcü Arı Evresi*/
                for j:=1 to D do
                Xo := Komsu_Cozumleri_Uret( Xkij )
                if (Xo, Xkij den daha kötü) ise
                Xkij don't change
                else if (Xo, Xkij den daha iyi) ise
                Xki := Xo
                end for

if (Xki, best_sol dan daha iyi) then
best_sol := Xki
end for
until (sonuç değeri uygun değil)

return best_sol
        end procedure
```

3.1.6.3. Gözcü arı evresi

Orijinal ABC algoritmasında belirtildiği gibi, gözcü arılar hangi işçi arının takip edileceğine karar vermek için işçi arıların sallanma dansını izlerler. Gözcü arı denklem 3.4'de hesaplanan p_i olasılık değerine bağlı x_i işçi arı tarafından bulunan bir yemek kaynağını tercih eder.

$$P_i = \frac{fit_i}{\sum_{i=1}^{SN} fit_i} \tag{3.4}$$

Buradaki fit_i değeri *i.* yemek kaynağının nektar miktarını ifade eder. Herhangi bir yemek kaynağına gidecek olan gözcü arı sayısı kaynaktaki nektar miktarına bağlı olarak değişir. Buradaki, fit_i değeri ne kadar yüksekse, *i.* yemek kaynağının seçilme olasılığı o kadar yüksek demektir.

Gözcü arı x_i isimli işçi arı tarafından bulunan yemek kaynağını seçer ve gözcü arı denklem 3.2 ve 3.3'e göre x_i konumumda bir modifikasyon yapar. İşçi arı evresinde olduğu gibi, modifiye edilmiş yemek kaynağı x_i 'den daha iyi nektara sahipse ya da nektar miktarları birbirine eşitse, modifiye edilmiş yeni yemek kaynağı x_i ile yer değiştirir ve arı popülasyonun yeni bireyi olur (Wei ve Hanning 2012).

3.1.6.4. Kâşif arı evresi

Bu evrede, yeni potansiyel yemek kaynakları kâşif arılar tarafından bulunur. Eğer bir yemek kaynağı daha önceden belirlenmiş "limit" adı verilen sayıdaki döngü içerisinde iyileştirilemezse, bu durumda yemek kaynağının terk edilmesi gerekir (Wei ve Hanning 2012).

3.2. Yerçekimsel Arama Algoritması

Fizikte yerçekimi kanununa göre, kütleye sahip her nesne birbirine doğru hızlanırlar. Yine Newton yerçekimi kanununda, kütleye sahip her bir nesne diğer nesneyi belli bir güçle çekmektedir. Bu kural *yerçekimsel güç* olarak adlandırılmaktadır. GSA, Newton yerçekimi ve hareket kanunlarından esinlenerek geliştirilmiş yeni bir optimizasyon algoritmasıdır. GSA'da kütleler olarak isimlendirilen bir dizi ajan Newton yerçekimi ve hareket kanunlarına göre modellenerek en uygun çözümü bulmak için kullanılır (Rashedi ve ark 2009, 2010).

GSA'yı başlatmak için i. kütlenin (ajan) konumunun denklem 3.5'deki gibi tanımlandığı *s* kümeli bir sistem varsayalım:

$$x_i = \left(x_i^1, \dots, x_i^d, \dots, x_i^n\right),\ i = 1,2, \dots, s \qquad (3.5)$$

Burada x_d^i, *d.* boyuttaki *i.* kütlenin konumudur, *n* ise arama uzayının boyutudur. Her bir ajanın kütlesi, uygunluk değeri ile temsil edilir. Ajana ait mevcut kütle

dizisindeki diğer ajanların uygunluğu denklem 3.6'da hesaplanır (Rashedi ve ark 2009, 2010).

$$q_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \tag{3.6}$$

M_i değeri denklem 3.7'de ifade edildiği gibi mevcut kütlenin toplam kütleye oranına eşittir.

$$M_i(t) = \frac{q_i(t)}{\sum_{j=1}^{x} q_j(t)} \tag{3.7}$$

Bu denklemlerde $M_i(t)$ ve $fit_i(t)$ sırasıyla t anındaki i. ajanın kütlesini ve uygunluk değerini ifade etmektedir. *best(t)* ve *worst(t)* ifadeleri t anında en düşük ve en yüksek değeri bulmak için denklem 3.8 ve 3.9'daki gibi tanımlanır (Rashedi ve ark 2009, 2010).

$$best(t) = \min(fit_j(t)) \; and \; j \in 1, \ldots, s \tag{3.8}$$

$$worst(t) = \max(fit_j(t)) \; and \; j \in 1, \ldots, s \tag{3.9}$$

Yerçekimi kanununa bağlı olarak bir ajanın ivme hesabı yapılırken, diğer kütleler (ajanlar) tarafından o ajan üzerine etki eden toplam güç hesaplanır. Söz konusu ajana uygulanan toplam güç miktarı denklem 3.10'a göre hesaplanmaktadır. Toplam güç hesaplandıktan sonra hareket kanununa göre ajanın ivmesi denklem 3.11'e göre hesaplanır. Ajana ait ivme değeri mevcut hızıyla toplanıp yeni hız vektörü denklem 3.12'ye göre hesaplanır. Hız vektörü bulunduktan sonra, ajanın yeni konumu denklem 3.13'e göre bulunmuş olur (Rashedi ve ark 2009, 2010).

$$F_i^d(t) = \sum_{j \in kbest \;\&\; j \neq 1} rand_j G(t) \frac{M_j(t) M_i(t)}{R_{ij}(t) + \varepsilon} \left(x_j^d(t) - x_i^d(t) \right) \tag{3.10}$$

$$a_i^d(t) = \frac{Fi^d(t)}{M_i(t)} = F_i^d(t) = \sum_{j \in kbest \;\&\; j \neq 1} rand_j G(t) \frac{M_j(t)}{R_{ij}(t) + \varepsilon} \left(x_j^d(t) - x_i^d(t) \right) \tag{3.11}$$

$$v_i^d(t+1) = rand_i \times {v_i}^d(t) + a_i^d(t) \tag{3.12}$$

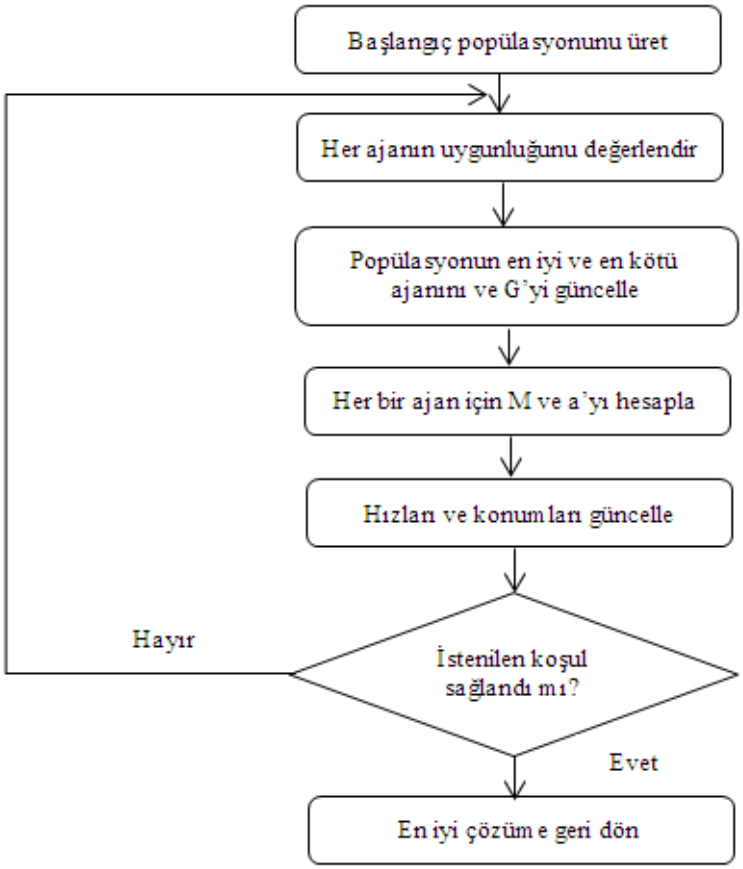
$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \tag{3.13}$$

Denklemlerde yer alan $rand_i$ ve $rand_j$ değerleri [0,1] aralığında dağıtılan rastgele iki sayıdır. ε küçük bir değer olup, bu değerini kullanmaktaki amaç denklemin tanımsızlığını ortadan kaldırmaktır. $R_{ij}(t)$, i ve j ajanları arasındaki öklid uzunluğunu ifade etmektedir ve denklem 3.14'deki gibi hesaplanır.

$$R_{ij}(t) = \| x_i(t), x_j(t) \|_2 \tag{3.14}$$

En iyi uygunluk değerini ifade eden k_{best}, en büyük kütleye sahip olan tüm ajanlar içindeki ilk k ajanlarının bir dizisini tanımlar. k_{best} başlangıçta k_0'da başlatılan ve zamanla azaltılan bir zaman fonksiyonunu ifade etmektedir. Buradaki k_0 değeri, s'nin (toplam ajan sayısı) kümesidir ve doğrusal olarak 1'e doğru azalır. GSA'da $G(t)$ yerçekimsel sabiti, β sabit katsayıyı, t şu anki iterasyon sayısını, t_{max} bitiş iterasyon sayısını, G_o ise bir başlangıç değerini temsil etmektedir (Rashedi ve ark 2009, 2010). Denklem 3.15'de G_o'a bağlı olarak her bir iterasyondaki $G(t)$ ifadesi verilmiştir. GSA algoritmasına ait akış diyagramı Şekil 3.2'de gösterilmiştir.

$$G(t) = G_0 e^{-B(\frac{t}{tmax})} \tag{3.15}$$



Şekil 3.2. GSA'nın akış diyagramı

3.2.1. İkili (binary) yerçekimsel arama algoritması

ÖS ve boyut azaltma (Bereta ve Burczyński 2007, Wang ve ark 2007, Chuang ve ark 2008, Zeng ve ark 2009, Pal ve Maiti 2010) veri madenciliği (Srinivasa ve ark 2007) ve hücre formasyonu (Wu ve ark 2008) gibi ikili vektör olarak çözümün kodlanmasının gerekli olduğu çok fazla optimizasyon problemi vardır. Buna ek olarak, gerçek uzayda tanımlı problemler de ikili uzayda düşünülebilir. Bu çözüm ikili modda birkaç biti olan gerçek rakamları göstermek içindir.

Bu bölümde, GSA'nın ikili versiyonu sunulacaktır. Bunu gerçekleştirmek için, GSA'nın bazı temel kavramları değiştirilecektir. Ayrık ikili uzayda, her boyut sadece 0 ya da 1 değeri alabilir. Bir boyuttaki hareket ilgili değişken değerinin 0'dan 1'e değişmesi ya da tam tersinin gerçekleşmesi anlamına gelir. GSA'nın ikili moduna başlamak için güç, ivme ve hız güncelleme prosedürlerinin sürekli algoritmalara benzer olduğu düşünülebilir. Sürekli (continuous) ve ikili (binary) GSA arasındaki temel fark; sürekli algoritmada pozisyon güncellemesi sürekli değer kümesi içinde tanımlanırken, ikili algoritmada pozisyon güncellemesi 0 ve 1 değerleri arasında tanımlanmaktadır. Bu değişim kütle hızına göre yapılması gerekir. Pozisyon güncellemesi kitle hızına göre hesaplanılmış bir olasılıkla mevcut bit değerlerini değiştirme mantığına dayanmaktadır. Başka bir deyişle, İGSA hızını denklem 3.12'ye göre günceller ve verilen olasılık dahilinde ya 0 ya da 1 olarak yeni pozisyonu atar.

Pozisyon güncelleme olasılığına göre hızı yerleştirecek (mapping) bir transfer fonksiyonu tanımlanmadan önce, GSA'nın bazı temel kavramlarını hatırlamak faydalı olacaktır. Buna göre;

- Büyük mutlak hız değeri kütlenin mevcut pozisyonunun hazır olmadığını ve büyük bir hareketin optimum pozisyona ulaşması için gerekli olduğunu gösterir.
- Küçük mutlak hız değeri kütlenin mevcut konumunun optimum pozisyona yakın olduğunu ve optimum pozisyona ulaşmak için çok küçük bir mesafe olduğunu işaret eder. Ardından, optimum çözümü bulduktan sonra, hız değeri 0 olur.

GSA'nın yukarıdaki kavramlarına dayanan İGSA'nın uygulanmasında, aşağıdaki kavramlar dikkate alınmalıdır.

- Büyük mutlak hız değeri önceki konumuna nazaran kütlenin pozisyonunu değiştirecek yüksek bir olasılık sağlamalıdır. (1'den 0'a veya tam tersi)
- Küçük mutlak hız değeri pozisyon değişiminde küçük bir olasılık sağlamalıdır. Başka bir deyişle, hızın 0 değeri kütle pozisyonun iyi olduğunu ve değiştirilmemesi gerektiğini gösterir.

Yukarıda bahsedilen kavramlara dayanan, uygun bir olasılık fonksiyonu tanımlanmalıdır. Öyle ki, küçük $\left|v_i^d\right|$ değeri için, x_i^d değişme olasılığı 0 değerine yakın olmalı ve büyük $\left|v_i^d\right|$ değeri için x_i^d değişme olasılığı yüksek olmalıdır. v_i^d yi olasılık fonksiyonuna göndermek için $S(v_i^d)$ fonksiyonu tanımlanır. $S(v_i^d)$ [0,1] aralığında sınırlandırılmış olması gerekir ve vi^d'nin artmasıyla artar. Denklem 3.16'da $S(v_i^d)$ fonksiyonu gösterilmiştir.

$$S\left(v_i^d(t)\right) = \left|\tanh(v_i^d(t))\right| \tag{3.16}$$

$S(v_i^d)$ is hesaplandığında, ajanlar denklem 3.17'de açıklanan kurala göre hareket edecektir.

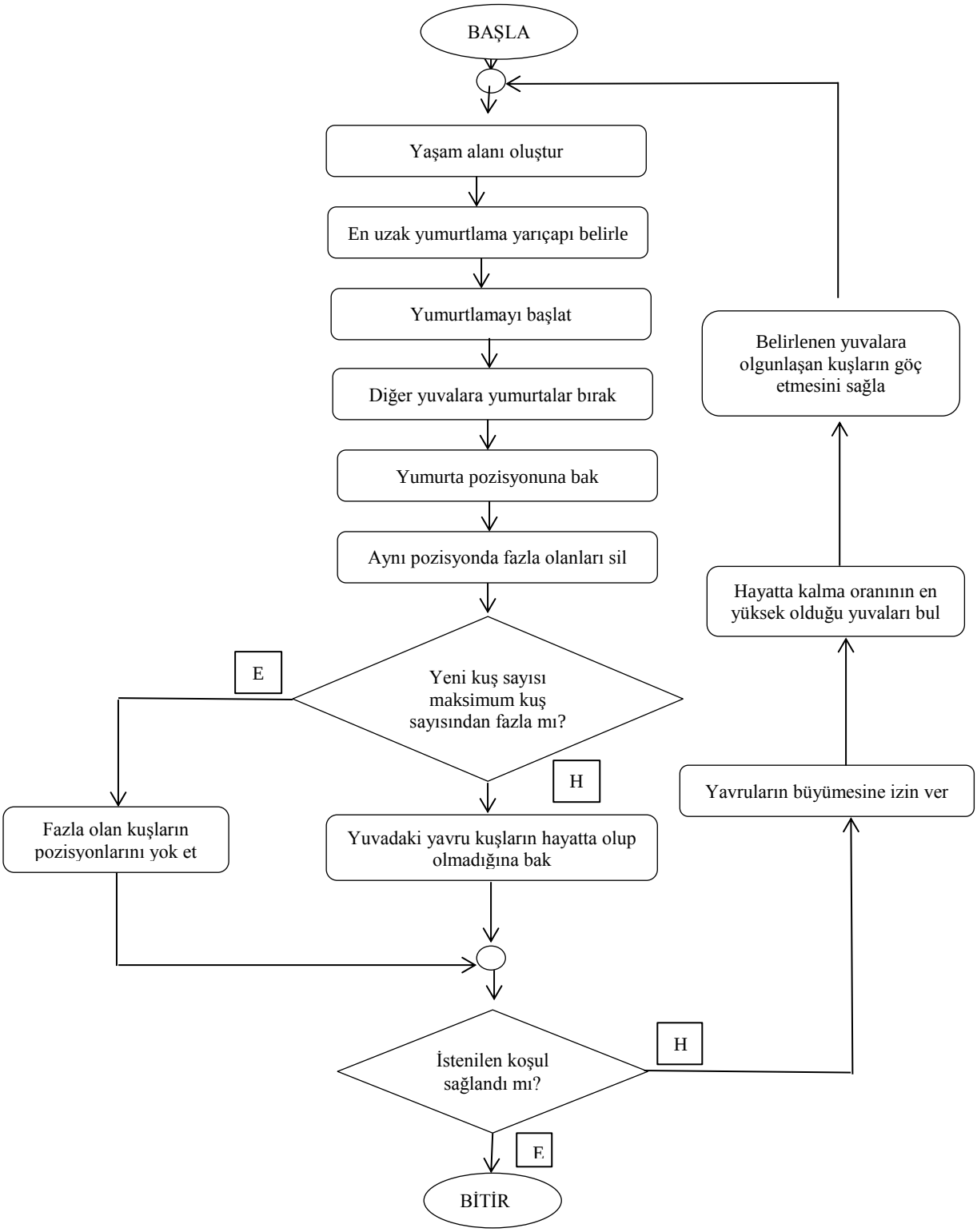
$$if\ rand < S\left(v_i^d(t+1)\right)\ then\ x_i^d(t+1) = complement\left(x_i^d(t)\right)$$
$$else\ x_i^d(t+1) = x_i^d(t) \tag{3.17}$$

İyi bir yakınsama oranını yakalamak için, hızın sınırlandırılması gerekir. Burada Hamming mesafesine dayanılarak hesaplanan *R* uzaklık değeri dikkate alınır (Rashedi ve ark 2010).

3.3. Guguk Kuşu Arama Algoritması

Guguk Kuşu arama (CSO) algoritması Yang ve Deb tarafından geliştirilmiş sürü zekasına dayanan yeni bir optimizasyon algoritmasıdır (Yang ve Deb 2009),(Inderscience 2010). Algoritmanın oluşturulmasında, doğadaki bazı guguk kuşu türlerinin yuvalarında bulunan yumurtaları başka bir yuvaya bırakan guguk kuşu yavrularının parazit davranışlarından ilham alınmıştır. Bazı ev sahibi kuşlar yuvaya izinsiz giren kuşlarla çatışmak zorunda kalabilirler. Örneğin, ev sahibi bir kuş kendisine ait olmayan yumurtaları bulması halinde ev sahibi kuşun yapacağı iki durum söz konusudur. Bunlarda ilki, yabancı yumurtaları yuvadan atmak; ikincisi ise, ev sahibi kuş

kendi yuvasını terk edip herhangi bir yere yeni bir yuva kurmaktır (Zheng ve ark 2012). Şekil 3.3'te CSO algoritmasının akış diyagramı verilmiştir.



Şekil 3.3. CSO algoritmasının akış diyagramı

Guguk kuşu araması bu tip üreme tavrına göre geliştirilmiş olmasından dolayı çeşitli optimizasyon problemleri için uygulanabilmektedir. CSO algoritması uygulamalarda diğer meta-sezgisel algoritmalardan daha iyi performans ortaya koyduğu görülmektedir.

CSO algoritmasında her bir yumurta bir çözümü temsil eder ve her bir guguk kuşu yumurtası yeni bir çözümü ifade eder. Burada amaç yuvadaki çok iyi olmayan mevcut çözümle yer değiştirecek potansiyel olarak daha iyi yeni çözümleri (guguk kuşları) kullanmaktadır. Algoritma, her bir yuvanın bir dizi çözümleri temsil eden çoklu yumurtalara sahip olduğu daha karışık durumlara uzanabilir. CSO, üç temel kurala dayanır (Yang ve Deb 2009).

1. Her bir guguk kuşu tarafından bir seferde bir yumurta verilir ve rastgele seçilen bir yuvaya guguk kuşunun yumurtası bırakılır.

2. Yumurta kalitesi en iyi olan yuva sonraki iterasyona taşınır.

3. Eldeki mevcut ev sahibi yuva sayısı sabittir ve bir kuş tarafından bırakılan yumurta $p_a \epsilon (0,1)$ ihtimaline göre ev sahibi kuş tarafından keşfedilir.

Diğer evrimsel hesaplamalar gibi, CSO da guguk kuşlarının başlangıç popülasyonu ile başlar. Başlangıç guguk kuşları kendi yuvalarına bıraktıkları yumurtaya sahiptirler. Ev sahibi kuşun yumurtasına daha çok benzerlik gösteren bu yumurtaların bazıları büyüme imkânına sahiptir ve bu yumurtalar aynı zamanda olgun bir kuşa dönüşürler. Diğer yumurtalar ise ev sahibi kuş tarafından bulunur ve yok edilirler. Gelişen bu yumurtalar o alanda bulunan yuvaların uygunluğunu ortaya çıkarır. O alanda ne kadar yumurta hayatta kalmışsa, o bölgede o kadar kazanç sağlanır. Bundan dolayı çoğu yumurtanın hayatta kaldığı bu pozisyonu CSO'nun iyileştireceği anlamına gelmektedir.

CSO algoritması yumurtalarının hayatta kalma oranını en yüksek oranda tutmak için yumurtalarını bırakacağı en uygun alanı araştırır. Geriye kalan yumurtaların büyüme ve olgun kuş haline dönüşmesi sürecinden sonra guguk kuşları birkaç toplum meydana getirirler. Her bir toplumun yaşamını devam ettirebileceği bir habitatı bulunmaktadır. Tüm gruplar için içindeki en iyi habitat diğer toplumlardaki guguk kuşları için gidilecek yeni bir hedef olacaktır. Hedef habitat belirlendikten sonra guguk kuşları en iyi habitata doğru göç ederler. Guguk kuşları en iyi habitatın etrafında herhangi bir yerde yaşayacaklardır. Her bir guguk kuşunun sahip olduğu yumurta sayısı

ve bir de guguk kuşunun en iyi habitata olan uzaklığı düşünülerek yumurtlama sahasındaki herhangi bir yumurta en iyi habitata gönderilir. Bu süreç, popülâsyon çoğunluğu aynı pozisyon etrafında toplanana kadar ve maksimum kazanç değerine sahip en iyi pozisyon elde edilene kadar devam eder (Zheng ve ark 2012).

3.3.1. Başlangıç popülasyonu oluşum evresi

Herhangi bir optimizasyon problemini çözmek için, problem değişkenlerinin değerleri bir dizi olarak oluşturulmalıdır. GA ve PSO'da bu diziler sırasıyla kromozom ve parçacık pozisyonu olarak adlandırılır. CSO'da ise diziler habitat olarak isimlendirilir. N boyutlu bir optimizasyon probleminde, bir habitat guguk kuşlarının şu anki mevcut pozisyonunu temsil eden $1xN_{var}$ isimli bir dizisidir. Bu dizi aşağıdaki gibi tanımlanır.

$$\text{habitat}=[x_1,x_2,\ldots,x_{Nvar}]$$

Her bir değişken değeri $(x1,x2,\ldots,x_{Nvar})$ float bir sayıdır. Bir habitatın kazancı $(x1,x2,\ldots,x_{Nvar})$ noktasındaki f_p fonksiyonunun kazanç hesabı aşağıdaki gibi elde edilir.

$$\text{Kazanç} = f_p(\text{habitat}) = f_p(x_1,x_2,\ldots,x_{Nvar})$$

Görüldüğü gibi CSO, kazanç fonksiyonunu maksimize eden bir algoritmadır. Zarar minimizasyon probleminde kullanmak için aşağıdaki kazanç fonksiyonu kolaylıkla maksimize edilebilir.

$$\text{Kazanç} = -\text{ Zarar(habitat)} = -\ f_c(x_1,x_2,\ldots,x_{Nvar})$$

Optimizasyon algoritmasına başlamak için, $N_{pop}\ x\ N_{var}$ boyutunda bir aday habitat matrisi oluşturulur. Başlangıçta guguk kuşu habitatlarının her biri için rastgele olarak üretilmiş birkaç yumurtanın var olduğu varsayılır. Doğada her bir guguk kuşu 5 ila 20 yumurta bırakır. Bu değerler farklı iterasyonlarda her bir guguk kuşu adına yumurtalama sayısının yüksek ve düşük limitleri olarak kullanılır. Gerçek guguk kuşlarının başka bir davranışı da yumurtalarını habitatlarından maksimum uzaklık içerisine bırakıyor olmalarıdır. Bundan sonra, bu maksimum aralık yumurta bırakma yarıçapı (YBY) olarak isimlendirilir. Yüksek limit ve düşük limit değişkenleri olan bir optimizasyon probleminde, her bir guguk kuşu toplam sayıda yumurta, şu anki guguk

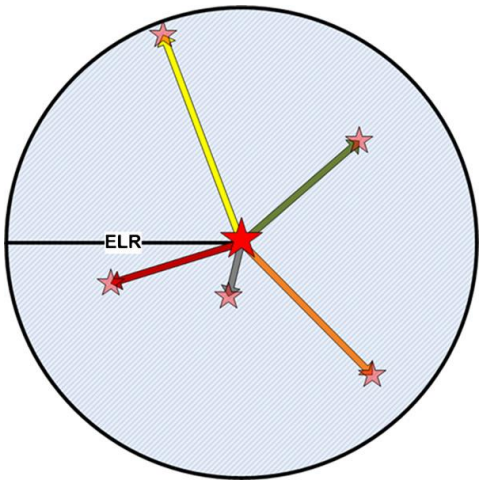
kuşunun yumurtaları ve bir de düşük ve yüksek değişken limitleriyle orantılı olan YBY'ye sahiptir. Bundan dolayı YBY denklem 3.18'deki gibi tanımlanır.

$$\text{YBY} = \alpha \times \frac{\text{Şu anki guguk kuşunun yumurta sayısı}}{\text{Toplam yumurta sayısı}} \times (var_{hi} - var_{low}) \quad (3.18)$$

α değeri YBY'nin maksimum değerini tutacağı varsayılan tamsayı bir değerdir. var_{hi} değeri yüksek limiti, var_{low} değeri düşük limiti ifade eder.

3.3.2. Guguk kuşunun yumurtlama evresi

Her bir guguk kuşu kendi YBY'si içerisine diğer bir kaç ev sahibi kuşun yuvalarındaki yumurtalarını rastgele bırakarak başlar. Şekil 3.4 bu kavrama daha net bir bakış sunar.



Şekil 3.4. YBY içerisine bırakılan rastgele yumurta yuvaları (Rajabioun 2011)

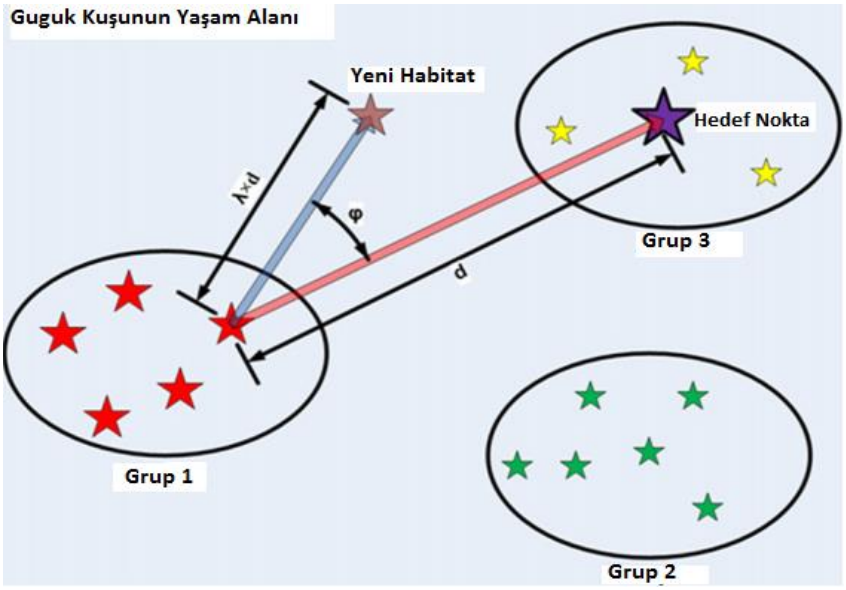
Şekil 3.4'de yumurta merkezi kırmızı yıldız, 5 yumurtası olan guguk kuşunun başlangıç habitatıdır. Pembe yıldız yumurtaların yeni yuvasını temsil etmektedir. Tüm guguk kuşlarının yumurtaları ev sahibi kuşların yuvalarına bırakıldıktan sonra, ev sahibi kuş tarafından saptanan yumurtaların bazıları ev sahibi kuşların kendi yumurtalarına daha az benzerler ve yuvanın dışına atılır. Bundan dolayı yumurta bırakma işlemi bittikten sonra, daha az kazanca sahip değeri olan öldürülür. Bu yumurtaların büyüme şansı yoktur. Kalan yumurtalar yuvalarında büyürler ve ev sahibi kuşlar tarafından beslenirler. Guguk kuşlarının yumurta bırakmaları ile ilgili ilginç olan başka bir nokta yuvadaki sadece tek bir yumurtanın büyüme şansına sahip olmasıdır. İki günden sonra

ev sahibi kuşun kendi yavruları açlıktan ölür ve sadece guguk kuşu yavrusu yuvada kalır (Rajabioun 2011).

3.3.3. Guguk kuşlarının göç etme evresi

Genç guguk kuşları büyüdüğünde ve olgun bir kuş olduğunda, kendi alanlarında ve toplumlarında ileride yaşarlar. Yumurta bırakma işlemi olacağı zaman, guguk kuşları ev sahibi kuşların yumurtalarına daha benzer yumurtaları olan ve yeni genç kuşlar için daha fazla yemek olan daha iyi habitatlara göç ederler. Guguk kuşu grupları farklı alanlarda oluşturulduktan sonra, en iyi kazanç değerine sahip toplum (grup) diğer guguk kuşlarının göç etmesi için hedef nokta olarak seçilir. Yetişkin guguk kuşları çevredeki her yerde yaşadıkları için hangi guguk kuşunun hangi gruba ait olduğunu bilmek zordur. Bu problemi çözmek için guguk kuşlarının gruplaması *k-means* kümeleme metoduna göre yapılır. Guguk kuşları grubu oluşturulduktan sonra ortalama kazanç değeri hesaplanır. Bundan sonra ortalama kazançların maksimum değeri hedef grubu belirler ve sonuç olarak grubun en iyi habitatı göç eden guguk kuşları için yeni hedef habitatıdır.

Guguk kuşları hedef noktaya doğru hareket ettiğinde, hedef habitata tüm yol boyunca uçmazlar. Onlar yolun bir kısmını uçarlar ve bir de sapma yaparlar. Bu hareket açıkça Şekil 3.5'de gösterilmektedir. Şekil 3.5'de gösterildiği gibi her bir guguk kuşu hedef habitata doğru tüm mesafenin sadece yüzde λ'lık kısmını uçar ve guguk kuşunun ϕ radyanlık bir sapması vardır. Bu iki parametre (λ *ve* ϕ) tüm çevrede çok daha fazla pozisyon aramasına yardımcı olur. Her bir guguk kuşu için λ ve ϕ denklem 3.19'deki gibi tanımlanır (Rajabioun 2011).



Şekil 3.5. Hedef noktaya doğru göç eden bir guguk kuşu örneği

$$\lambda \sim U(0,1)$$
$$\varphi \sim U(-\omega, \omega) \tag{3.19}$$

$\lambda \sim U(0,1)$ ifadesi 0 ila 1 arasında homojen olarak dağıtılmış random bir sayı anlamına gelmektedir. ω ifadesi, hedef habitattan sapmayı sınırlandıran bir parametredir. $\pi/6$'lık bir ω, guguk kuşu popülasyonun global maksimum kazanca iyi bir yakınsama yapması için yeterli olduğu görülmektedir. Tüm kuşlar hedef noktaya göç ettiklerinde ve yeni habitatlar belirlendiğinde, bir kaç yumurta her bir olgun guguk kuşuna verilir. Her bir kuşa ait yumurta sayısı düşünülerek her bir guguk kuşu için YBY hesaplanır. Bundan sonra yeni yumurta bırakma işlemi tekrar başlatılır.

3.3.4. En kötü habitattaki guguk kuşlarını yok etme evresi

Kuş popülasyonunda sürekli bir denge olduğu gerçeğinden dolayı N_{max} ifadesi çevredeki canlı guguk kuşlarının maksimum sayısını kontrol eder ve sınırlandırır. Bu denge yemek sınırlandırmasından ve aynı zamanda yumurtalar için hazır yuvayı bulmadaki yetersizlikten dolayıdır. Sadece daha iyi kazanç değerine sahip N_{max} sayıdaki guguk kuşu hayatta kalır, diğer guguk kuşları habitattan çıkartılır.

3.3.5. Yakınsama evresi

Birkaç iterasyondan sonra, tüm popülasyon ev sahibi kuşlara maksimum benzerliği olan yumurtaya sahip olan ve aynı zamanda maksimum yemek kaynakları olan en iyi habitata hareket eder. Bu habitat maksimum kazancı şimdiye kadar üretir. En iyi habitatta en az yumurta kaybı olacaktır. Aynı habitata tüm guguk kuşlarının % 95'den fazlasının yakınsaması CSO'yu sonlandırır. CSO'nun temel adımları Şekil 3.3'te gösterilmiştir (Rajabioun 2011).

3.3.6. İkili guguk kuşu arama algoritması

Çoğu problem ayrık bir yapıya sahiptir. Buna ek olarak ayrık işlemler bağlantılı problemlerin çoğunu da çözmek için kullanılır. Bundan dolayı ayrık CSO algoritmasına ihtiyaç duyulur. Çünkü hem ayrık hem bağlantılı problemler aynı ayrık alanda çözülebilir. Guguk kuşu aramanın temel versiyonu şu zamana kadar farklı problemleri çözmek için kullanılmıştır fakat bu model habitatlara doğru gelen guguk kuşlarının hareketindeki bazı değişimleri yaparak daha iyi ayrık bir modele dönüştürülebilir (Shirin ve ark 2012).

3.3.6.1. Guguk kuşlarının hareketindeki seviyeyi değiştirme

Tüm guguk kuşlarının her bir algoritma tekrarında daha iyi habitata ulaşmak için en iyisini yaptığı düşünüldüğünde, habitatların bazıları en iyisi olacaktır. Böylelikle klasik guguk kuşu modeli, GA'daki kromozom yapıları ile tamamen aynı olan her bir habitatın yapısını tanımlayarak ayrık bir modele dönüştürebilir ve habitata doğru guguk kuşlarının göç işlemi çaprazlama işlemi eklenerek kullanılabilir.

Kromozom olarak habitat yapıları tanımlandıktan sonra, uygun çaprazlamanın seçilmesi gerekir. 0 ve 1 ifadeleri Graph Bisection probleminde her bir yuvadaki yumurtaların sırasıyla varlığını ya da yokluğunu temsil eder. 0 ifadesi yumurtanın yok olduğu anlamına gelir, 1 ifade ise yuvada yumurtanın olduğu anlamına gelir.

Yarı düzenli çaprazlama göç (en iyi habitata doğru guguk kuşlarının hareketi) için kullanılır. İlk olarak yuvaların yüzde 50'si yarı düzenli çaprazlamada her bir habitattan rastgele seçilir. Sonra en iyi habitat aynı yerdeki habitatla yer değiştirir.

Global optimum cevaba doğru hızlı yakınsama, lokal optimum noktalarda zorluğa ilişkin düşük olasılık ve hazır olmayan cevabı yok etme sayesinde daha iyi

noktalara doğru popülasyonun genel hareketi, ilgili algoritmanın bazı avantajları olarak söylenebilir. Çizelge 3.3 ve 3.4'te guguk kuşlarının ikili hareketlerine ilişkin örnekler verilmiştir (Shirin ve ark 2012).

Çizelge 3.3. Yuvanın rastgele seçimi için bir örnek (Shirin ve ark 2012)

Best habitat		**1**	**0**	**1**	**1**	**1**	**1**
Habitat1		0	1	0	1	0	0
Habitat2		0	0	0	1	0	0
……							
Habitat n		1	1	0	0	1	0

Çizelge 3.4. En iyi habitata doğru guguk kuşlarının hareketi (Shirin ve ark 2012)

Best habitat		**1**	**0**	**1**	**1**	**1**	**1**
Habitat1		1	1	1	1	0	0
Habitat2		1	0	0	1	1	0
……							
Habitat n		1	0	0	0	1	1

3.4. Yarasa Algoritması

Yarasa Algoritması, yarasaların *ekolokasyon* adı verilen sesin yankılanmasından faydalanılarak bir cismin bulunduğu yön ve uzaklık saptama davranışından esinlenerek Yang tarafından sürü zekasına dayanan yeni bir optimizasyon algoritması olarak önerilmiştir (Yang 2010).

Yarasalar da dâhil ekolokasyon kullanan tüm canlılar (yunuslar, balinalar, kır fareleri ve bazı kuş türleri) belli bir frekansta sinyal yayarlar. Bu sinyaller çoğunlukla insanın algılayabileceği (yaklaşık 20 kHz) ses sınırlarının ötesindedir. Yarasalar, avlarının konumlarını belirlemede, aralarındaki iletişimi sağlamada, tamamen karanlık ortamlarda dahi her türlü nesneyi algılayıp onlara çarpmadan hareket etmelerinde ve hareket halindeki herhangi bir böceği ayırt etmede ekolokasyonu kullanırlar (Yang 2010).

Yang (2010) yarasaların ekolokasyon karakteristiklerinden bazılarını ideal hale getirmiş ve aşağıdaki kurallara göre algoritmanın oluşumunu temellendirmiştir (Yang 2010).

1. Yarasalar mesafeyi algılamak için ekolokasyon kullanırlar. Hatta yarasalar kendi avı ile arka plandaki engeller arasındaki mesafeyi olağanüstü özellikleriyle bilirler.

2. Yarasalar avını aramak için x_1 pozisyonunda v_1 hızıyla, sabit f_{min} frekansı, değişken dalga boyu λ, ve ses şiddeti A_0 ile rastsal bir şekilde uçarlar. Yarasalar yaydıkları dalga boyunu ve sinyal emisyonunu $r \in [0,1]$ hedef bölgeye yakınlığına bağlı olarak otomatik bir şekilde ayarlar.

3. Ses şiddeti farkı şekillerde değişmesine rağmen, ses şiddetinin büyük değer A_0'dan, minimum sabit değeri A_{min} aralığında değiştiği varsayılır.

Genelde $[f_{min}, f_{max}]$ aralığındaki f frekansı $[mix, max]$ aralığındaki dalga boyuna karşılık gelmektedir. Örneğin [20 kHz, 500 kHz] aralığındaki bir frekans 0.7 mm'den 17 mm aralığındaki dalga boyuna karşılık gelmektedir. Ayrıca dalga boyu sabitlenirken frekansların değiştirilebileceği de Yang (2010) tarafından önerilmiştir. Basitçe, f'in $[0, f_{max}]$ aralığında olduğu varsayılır. Yüksek frekansların kısa dalga boyuna ve kısa mesafe gezilerine sahip olduğu bilinmektedir. Yarasalar için tipik aralık birkaç metredir. Sinyal oranı [0,1] aralığındadır; 0 hiç sinyal olmadığı, 1 ise sinyal emisyon oranının maksimum olduğu anlamına gelmektedir (Yang 2012).

Original BA Çizelge 3.5'de gösterilmektedir ve bileşenleri aşağıda açıklanmıştır;

- Başlangıç evresi (satır 2-4),
- Yeni çözümlerin oluşturulması (satır 6-7),
- Lokal arama evresi (satır 8-11),
- Rastsal uçuş yaparak yeni çözümlerin oluşturulması (satır 12-16),
- En iyi çözümün bulunması (Fister ve ark 2013)

Çizelge 3.5. BA'nın prosedürel kodu

1: Hedef fonksiyonu f(x), x=(x_1, …, x_d)^T
2: Yarasaları başlat x_i and v_i for $i = 1 \ldots n$
3: Frekans belirle $Q_i \in [Q_{min}, Q_{max}]$
4: Artış oranını belirle r_i and the loudness A_i
5: **while** $(t < T_{\max})$ // iterasyon sayısı
6: Frekansı ayarlayarak yeni çözümler üret, ve
7: Çözümleri ve hızı güncelle $[Eq.(2)\ to\ (4)]$
8: **if** $(rand(0,1) > r_i$
9: Yeni çözümler hesapla $NewBat = BestBat + Velocity$
10: En iyi çözüm etrafındanda yeni bir çözüm üret
11: **end if**
12: Rastgele uçarak yeni çözüm üret
13: **if** $(rand(0,1) < A_i\ and\ f(x_i) < f(x))$
14: Yeni çözümleri kaydet
15: r_i ve A_i değerini azalt
16: **end if**
17: Yarasaları sırala ve en iyi çözümü bul
18: **end while**

D-boyutlu bir arama uzayı için yarasaların v_i hızlarının ve x_i pozisyonlarının nasıl güncellendiği ve t anındaki yeni çözümler x_i^t ve hızlar v_i^t verilen denklemlerde gösterilmektedir.

$$f_i = f_{min} + (f_{max} - f_{min})\beta \tag{3.20}$$

$$v_i^t = v_i^{t-1} + (x_i^t - x_*)f_i \tag{3.21}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{3.22}$$

$\beta \in [0,1]$ normal dağılım ile belirlenen rastgele bir vektördür. x_* tüm yarasalar arasında tüm çözümlerin karşılaştırılması ile belirlenen o an ki en iyi çözümdür. $f_i\lambda_i$ hız artışı, ilgili problemin tipine bağlı olarak $f_i(veya\ \lambda_i)$ ikisinden biri sabitken, diğeri hız değişimlerini ayarlamak için kullanılabilir. Başlangıçta, her yarasaya $[f_{min}, f_{max}]$ aralığında rastgele bir frekans değeri atanır.

Lokal arama aşaması için, o anki en iyi çözümler arasından bir çözüm seçilir, bir yerel rastgele yürüyüş kullanılarak her yarasa için yeni bir çözüm üretilir. (Yang ve Gandomi 2012).

$$x_{new} = x_{old} + \epsilon A^t \tag{3.23}$$

$\varepsilon \in [-1,1]$ aralığında rastgele bir sayıyı, $A^t =< A_i^t >$ bu tüm yarasaların ortalama ses şiddetini temsil eder.

3.4.1. İkili Yarasa Algoritması

İkili arama uzayı bir hiper küp olarak düşünülebilir. İkili optimizasyon algoritmasının arama parçacıkları farklı sayıdaki bitlerin hareket ettirilerek sadece hiperküpün daha yakın ya da daha uzak köşelerine kaydırılabilir (Kennedy ve Eberhart 1997). Bundan dolayı BA'nın ikili versiyonunu tasarlamak için, hız ve pozisyon güncelleme işlemlerine ilişkin temel kavramların modifiye edilmesi gerekmektedir.

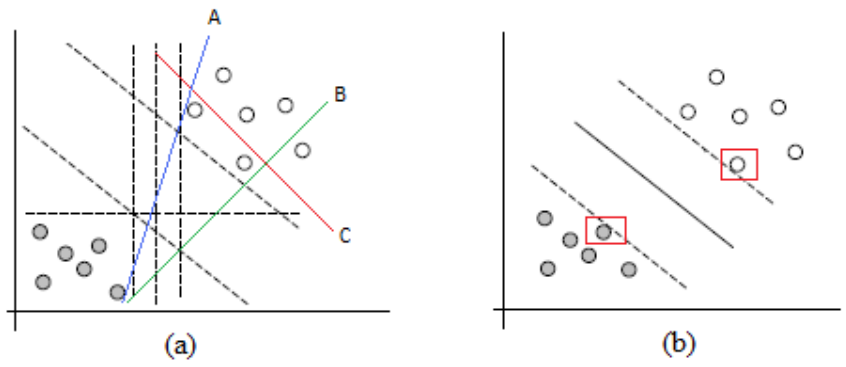
BA'nın sürekli versiyonunda yapay yarasalar sürekli değere sahip uzay içerisinde hız ve pozisyon vektörlerini kullanarak arama uzayında hareket edebilirler. Sonuç olarak, pozisyon güncelleme işlemi denklem 3.20'deki gibi pozisyonlara hız değeri eklenerek kolayca uygulanabilir. Fakat ayrık ikili uzayda pozisyon güncelleme işlemi farklıdır (Tasgetiren ve ark 2007). İkili bir uzayda, güncelleme işlemi sadece 0 ve 1 sayıları ile gerçekleştirileceğinden denklem 3.22 kullanılarak yapılamaz. Dolayısıyla, yarasaların pozisyonlarını 0'dan 1'e ya da tam tersi şekilde değiştirmek için hız değeri hesabı yapacak bir yöntemin geliştirilmesine ihtiyaç duyulmaktadır.

Ayrık ikili uzayda, pozisyon güncelleme 0 ve 1 değerleri arasındaki değişim anlamına gelmektedir. Bu değişimin yarasaların hızlarına göre yapılması gerekir (Mirjalili ve ark 2013).

4. DESTEK VEKTÖR MAKİNALARI

Destek vektör makineleri (DVM), doğrusal ve doğrusal olmayan verilerin sınıflandırılmasında kullanılan denetimli bir yöntemdir. Destek vektör makineleri metin sınıflandırma, parçacık tanıma ve yüz tanıma gibi pek çok gerçek dünya probleminde başarılı bir şekilde uygulanmaktadır. Destek vektör makineleri yönteminde temel düşünce pozitif ve negatif örnekleri ayıran en uygun düzlemin bulunmasıdır (Kumar ve Gopal 2009).

DVM'nin yüksek genelleme yeteneği bu yöntemin yüksek boyutlu veriler üzerinde uygulamasını mümkün kılmaktadır. DVM metin sınıflandırma konusunda birçok sınıflandırma yönteminden daha başarılı sonuç vermekle birlikte ÖS için de kullanılabilmektedir (Shima ve ark 2004). Destek vektör makineleri yüksek boyutlu özellik uzayı için etkili bir öğrenme başarısı elde etmektedir (Cachopo ve Oliveira 2006). Destek vektör makineleri eğitim verisini doğrusal olmayan bir dönüşümle yüksek boyuta çıkarmakta ve bu yeni boyut içinde en uygun ayırıcı düzlemi aramaktadır. En uygun düzlem iki sınıf arasındaki, sınıfların sınır noktaları arasındaki uzaklığı en yükseğe çıkaran düzlemdir. En uygun düzlemi bulmak için bu düzleme paralel ve sınırlarını oluşturacak iki düzlemin belirlenmesi gerekir. Bu düzlemleri oluşturan noktalar destek vektörü ismini almaktadırlar (Han ve Kamber 2006). Şekil 4.1'de sınıfları ayıran düzlem ve destek vektörler gösterilmektedir (Durmaz 2011).



Şekil 4.1. Sınıfları ayıran düzlemin belirlenmesi (a) En uygun düzlemin bulunması (b) Destek vektörler

Şekil 4.1'de gösterildiği gibi iki sınıflı uzayı B ve C düzlemleri tam olarak ayıramamaktadır. A düzlemi ise bu sınıfları en küçük aralıkla ayırmaktadır. Destek vektör makinelerinde hedeflenen nokta en büyük aralığı veren düzlemin seçilmesidir. Doğrusal veriler bu şekilde ayırmak mümkün iken doğrusal olmayan veriler için verilerin yüksek boyutta ifade edilmesi gerekir.

4.1. Çekirdek (Kernel) Fonksiyonları

Destek vektör makineleri ile matematiksel olarak $K(x_i,x_j) = \phi(x).\phi\ (x_j)$ şeklinde ifade edilen bir çekirdek fonksiyonu sayesinde doğrusal olmayan dönüşümler yapılabilmektedir. Bu işlem yapılarak verilerin yüksek boyutta doğrusal olarak ayrılmasını sağlamaktadır. Sonuç olarak, çekirdek fonksiyonu yardımıyla doğrusal olarak ayrılamayan iki sınıflı bir problemin çözümü ile ilgili karar kuralı aşağıdaki şekilde yazılabilir (Osuna ve ark 1997):

$$f(x) = sign(\sum_i \alpha_i y_i \varphi(x).\varphi(x_i) + b) \quad (4.1)$$

DVM ile gerçekleştirilecek bir sınıflandırma işlemi için öncelikle kullanılacak çekirdek fonksiyonu ve bu fonksiyona ait optimum parametrelerin belirlenmesi gereklidir. Literatürde çekirdek fonksiyonu olarak en sık kullanılan polinom, radyal tabanlı fonksiyon, Pearson VII (PUK) fonksiyonu ve normalleştirilmiş polinom çekirdekleri Çizelge 4.1'de formülleri ve parametreleriyle birlikte verilmiştir. Çizelge 4.1'de görüldüğü gibi her çekirdek fonksiyonu için bazı parametrelerin kullanıcı tarafından belirlenmesi gerekir. PUK çekirdeği için belirlenmesi gereken parametre sayısı iki iken, diğer fonksiyonlar için sınıflandırmayı gerçekleştirecek model oluşumunda bir parametrenin belirlenmesini gerektirmektedir.

Çekirdek fonksiyonları karşılaştırıldığında polinom ve radyal tabanlı çekirdeklerin daha sade ve anlaşılabilir olduğu görülmektedir. Polinomun derecesindeki artış matematiksel olarak basit görünse de, algoritmanın karmaşık bir hal almasına sebep olmaktadır. Bu da hem işlem süresini önemli ölçüde artırmakta hem de bir noktadan sonra sınıflandırma doğruluğunu azaltmaktadır. Buna karşın radyal tabanlı fonksiyonun çekirdek boyutu (γ) olarak ifade edilen parametresindeki değişimlerin sınıflandırma başarısına etkisinin daha az olduğu görülmüştür (Hsu ve ark 2010). Normalleştirilmiş polinom fonksiyonu ise veri kümesinin normalleştirilmesi yerine polinom çekirdeğine ait matematiksel ifadenin normalleştirilmesi amacıyla Arnulf ve ark (2001) tarafından önerilmiştir. Normalleştirilmiş polinom kernelinin polinom kernelinin genelleştirilmiş bir tipi olduğu söylenebilir. Diğer taraftan, PUK kerneli Pearson genişliği olarak bilinen (σ, ω) iki parametresi ile diğer çekirdek fonksiyonlarına göre daha karmaşık bir yapıya sahiptir. Bu iki parametre sınıflandırma doğruluğuna etki etmekte ve hangi parametre ikilisinin en iyi sonuç vereceği önceden bilinmemektedir.

Bu nedenle PUK kernelinin kullanımında en uygun parametre ikilisinin belirlenmesi önemli bir aşamadır (Kavzoğlu ve Çölkesen 2010).

Çizelge 4.1. Destek vektör makinelerinde kullanılan temel kernel fonksiyonları ve parametreleri

Kernel Fonksiyonu	Matematiksel İfadesi	Parametre
Polinom Kerneli	$K(x.y) = ((x.y)+1)^d$	Polinom derecesi (d)
Normalleştirilmiş Polinom Kerneli	$K(x.y) = \frac{((x.y)+1)^d}{\sqrt{((x.x)+1)^d.((y.y)+1)d}}$	Polinom derecesi (d)
Radyal Tabanlı Fonksiyon Kerneli	$K(x.y) = e^{-\gamma\lvert(x-x_i)\rvert^2}$	Kernel boyutu (γ)
Pearson VII (PUK) Kerneli	$K(x.y) = \frac{1}{\left[1+\left(\frac{2\sqrt{\lVert x-y\rVert^2\sqrt{2^{(1/\omega)}-1}}}{\sigma}\right)^2\right]^{\omega}}$	Pearson genişliği parametreleri (σ ve ω)

Kernel fonksiyonuna özgü parametrelerin yanında tüm destek vektör makineleri için düzenleme parametresi C'nin kullanıcı tarafından belirlenmesi gerekir. Bu parametre için olması gerekenden çok küçük veya çok büyük değerler seçilmesi durumunda optimum hiper-düzlem doğru belirlenemeyeceğinden sınıflandırma doğruluğunda ciddi düşüş beklenir. Diğer taraftan C = ∞ olması durumunda DVM modeli sadece doğrusal olarak ayrılabilen veri setleri için uygun hale gelir. Buradan da görüleceği üzere parametreler için uygun değerlerin seçimi DVM sınıflandırıcısının performansını direkt olarak etkileyen bir faktör durumundadır. Genellikle deneme ve hata stratejisi kullanılmasına karşın, çapraz doğrulama yaklaşımı başarılı sonuçlara ulaşılmasına olanak sağlamaktadır. Çapraz doğrulama yaklaşımında amaç oluşturulan sınıflandırma modelinin performansının belirlenmesidir. Bu amaçla veri kümesi iki kısma ayrılır. Birinci kısım sınıflandırmaya esas olan model oluşumunda eğitim verisi olarak kullanılırken ikinci kısım modelin performansının belirlenmesi amacıyla test verisi olarak işleme konur. Eğitim seti ile oluşturulan modelin test veri kümesine uygulanması sonucunda doğru sınıflandırılan örneklerin sayısı sınıflandırıcının performansını gösterir. Dolayısıyla çapraz geçerlilik yöntemi kullanılarak en iyi sınıflandırma performansının elde edildiği çekirdek parametrelerinin belirlenmesi ve sınıflandırmayı gerçekleştirecek modelin oluşturulması gereklidir (Kavzoğlu ve Çölkesen 2010).

5. ÖZELLİK SEÇİMİ VE ÖNERİLEN YÖNTEM

5.1. Özellik Seçimi

Sınıflandırma problemlerinde ve birçok öğrenme algoritmalarında önemli bir yere sahip olan ÖS, konu ile ilgili olan, en faydalı ve en önemli özelliklerin seçilerek veri kümesiyle ilgili özellik sayısının azaltılması için kullanılmaktadır. Bunun amacı, hesaplama yükünü azaltırken başarı oranını yükseltmektir.

5.1.1. Özellik Seçimi Tanımları

Özellik, nitelik veya değişken kavramları verinin görünümünden bahseder. Genellikle veriyi toplamadan önce, özellikler seçilir veya belirlenir. Özellikler ayrık, sürekli veya nominal (sözde) olabilir. Özellikler genelde aşağıdaki gibi karakterize edilir.

5.1.1.1. Relevant (Alakalı)

Alakalı özellikler çıkış üzerinde doğrudan etkisi olan özelliklerdir. Alakalı olan özelliklerin rolü gereksiz olarak düşünülemez.

5.1.1.2. Irrelevant (Alakasız)

Konu ile ilgisiz olan özellikler, çıkış üzerinde herhangi bir etkiye sahip olmayan özellikler olarak tanımlanırlar. Bunların değerleri her bir örneklemde rastgele üretilir.

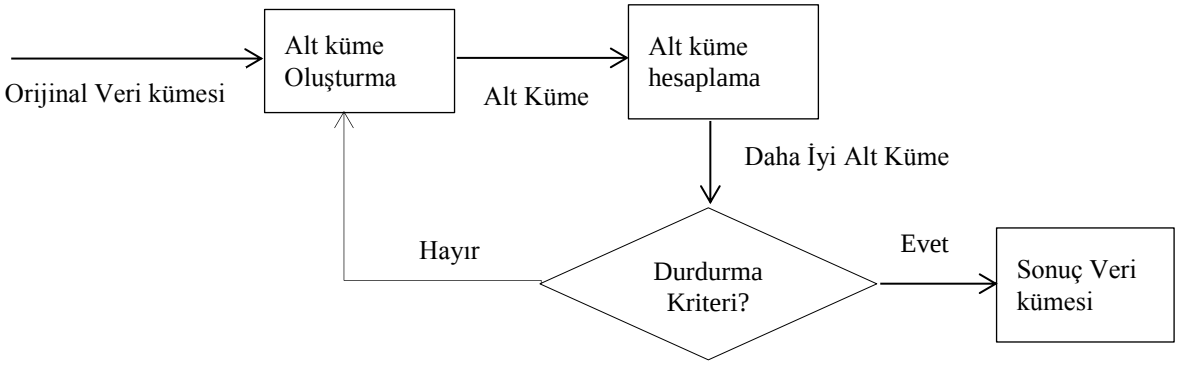
5.1.1.3. Redundant (Gereksiz)

Bir özelliğin diğerinin yerini alabildiği herhangi bir durumda ihtiyaç fazlalığının mevcut olduğu kabul edilir.

ÖS, kümeleme yapılırken tüm özellikler gerekli olmadığı için veri kümesini oluşturan tüm özellikler içerisinden en iyi özelliği/özellikleri seçme işlemidir. Bazı özellikler, çıkış değerlerini elde etmede gereksiz veya ilgisiz olabileceğinden öğrenme işlemine herhangi bir katkı sağlamıyor olabilir.

Alt küme seçimi olarak bilinen ÖS makine öğrenmesinde yaygın olarak kullanılan bir yöntemdir. ÖS işleminde, veri kümesinden elde edilen özellik alt kümesi, öğrenme algoritmasının uygulaması için seçilir. En iyi alt küme, çözüm uzayı için en yüksek doğruluk oranını verecek olan en küçük boyutlu veri kümesinden oluşur. Veri

kümesindeki geriye kalan önemsiz nitelikler yok sayılır. Bu işlem, önemli bir veri ön işleme aşamasıdır. ÖS'nin temel amacı, orijinal özelliklerin tamamını kullanmadan en yüksek oranda veri bütünlüğünü korumak ve aynı zamanda minimum özellik alt kümesini belirlemektir. Birçok gerçek dünya probleminde, ÖS gereksiz, yanıltıcı veya gürültülü verinin çokluğundan dolayı bir zorunluluk olarak düşünülmektedir. ÖS'den tamamıyla emin olabilmek için, tüm özellik kombinasyonlarının (alt kümelerinin) test edilmesi gerekmektedir. ÖS akış diyagramı Şekil 5.1'de gösterilmiştir.



Şekil 5.1. ÖS akış diyagramı

5.2. Özellik Seçiminin Avantajları ve Dezavantajları

ÖS'nin avantajları aşağıdaki gibi özetlenebilir;

- Gereksiz, sonuç değerine etkisi olmayan veya gürültülü verileri veri kümesinden kaldırmak,
- Veri analiz işlemlerinde öğrenme algoritmasının çalışma süresini hızlandırmak,
- Veri kalitesini iyileştirmek,
- Sonuç modeline ilişkin doğruluk oranını artırmak,
- Tahmin işleminde performansı yükseltmek,
- Verinin basit bir şekilde temsilini ve sunumunu sağlamak,
- Özellik uzay boyutunu düşürmek ve dolayısıyla depolama gereksinimlerini azaltarak algoritma hızını artırmaktır.

ÖS'nin dezavantajları ise aşağıdaki gibi belirtilebilir;

- Öğrenme (ÖS) için ek zamana ihtiyaç duymaktadır.

- İlgili özellik alt küme oluşturma işlemi, modellemede ek katmana ihtiyaç duymaktadır.

5.3. Özellik Seçim Algoritması

Özellik seçim algoritması hesaplamalı bir çözümdür. ÖSA'nın amacı ilgili çözüme ilişkin olan özellikleri belirlemektir. ÖSA'yı tanımlamak için literatürde birkaç konu yer almaktadır.

5.3.1. Arama organizasyonu

Hipotez uzayı ile ilgili genel strateji araştırılır. Bilindiği gibi *N tane özelliğe sahip D isminde bir veri kümesi için, D veri kümesinin* 2^N *adet aday alt kümesi vardır. N sayısı* değiştirilse bile, arama uzayı üstel olarak artacaktır. Bu da çok yoğun bir işlem hacmi gerektirdiğinden bu şekilde işleme tabi tutulmaması gerekmektedir. Bundan dolayı literatürde sıralı arama, üstel arama ve rastgele arama olarak bilinen stratejiler önerilmiştir.

5.3.1.1. Sıralı arama

Sıralı aramada, arama kriteri kendinden sonra gelecekler arasından sadece birini seçer. Bu sıralama iteratif olarak yürütülür ve muhtemel adım sayısı $O(n)$'dir. Sıralı arama bütünlük sağlar fakat optimal özellik, alt küme sonucunu vermez.

5.3.1.2 Üstel arama

Çok işlemli arama en iyi çözümü garanti eden optimal bir aramadır. Fakat, optimal çözümünün çok işlemli olmaması gereklidir. Farklı sezgisel fonksiyonlar optimal çözümden uzaklaşmadan arama uzayını azaltmak için kullanılabilir. Branch And Bound (Narendra ve Fukunaga, 1997) ve Beam Search (Doak, 1992) algoritmaları optimal çözümler için daha az sayıda çözümlerle hesaplanılır. Arama uzayının adım sayısı $O(n^2)$'dir.

5.3.1.3 Rastele arama

Rastele arama, rastgele seçilmiş alt küme ile başlar. Optimal alt küme çözümünü elde etmenin iki yolu vardır. Bunlardan biri, Las Vegas algoritması olarak bilinen tamamen rastgele bir yöntemdir (Brassard ve Bratley, 1996). Diğeri ise, yukarıdaki

sıralı yaklaşımdaki rastgeleliği içeren sıralı aramadır. Rastgelelik kavramı arama uzayındaki lokal optimumlardan kaçınmak içindir. Arama uzayı karmaşıklığı $O(n^2)$'dir.

5.3.2 Sonraki jenerasyon

Sonraki nesil için beş tane operatör vardır. Bunlar; ileri (forward), geri (backward), rastgele (random), bileşik (compound) ve ağırlık (weighting) olarak adlandırılmıştır. İleri operatöründe başlangıçta özellik alt kümesi boşken sınıflandırma başarısını artıran özellik her bir adımda seçilerek işlem tamamlanır. Geri operatöründe ise başlangıçta tüm özellikler seçili durumdadır. Her bir adımda sınıflandırma başarısını artıran özellik alt kümeden çıkartılarak işlem tamamlanır. Rastgele operatöründe işlemler rastsal olarak devam ettirilir. Bileşik operatöründe ileri ve geri operatörleri sınıflandırma başarısını artıracak şekilde kullanılır. Ağırlık operatöründe ise her bir özelliğin sınıflandırmaya katkısı ölçülerek işlemler yapılır.

5.4. Özellik Seçimi İçin Önerilen Sistem

Bu çalışmada ÖS'de en iyi alt kümeyi bulmak için dört farklı optimizasyon algoritması kullanılmıştır. Bu algoritmalar yukarıda belirtildiği gibi *ABC, CSO, BA ve GSA* algoritmalarının ikili versiyonlarıdır. Bu algoritmalara ÖS'yi gerçekleştirmek için veri setleri üzerinde çalışmadan önce ÖS için gerekli olmayan özellikler veri kümesinden çıkartılmıştır. Bununla birlikte ÖS sadece giriş değerleri üzerinde yapıldığından sınıf bilgisi de veri kümesinden çıkartılmıştır. Veri ön işlemeden sonra her bir veri kümesinin boyutuna göre başlangıç popülasyonundaki bireylerin her bir boyutuna [0,1] değerleri rastgele atanarak popülasyon oluşturulmuştur. Çözüm uzayının [0,1] olarak seçilmesindeki temel sebep, özellik seçim uygulamasında ilgili özelliğin çözüm uzayına eklenmesi ya da eklenmemesinin söz konusu olmasıdır. Popülasyonda yer alan "0" ifadesi ona karşılık gelen özelliğin seçilmediğini, "1" ifadesi ise ona karşılık gelen özelliğin yeni oluşturulacak veri kümesinde kullanılmak üzere seçildiğini ifade etmektedir. Özellik seçim uygulaması ayrık bir sistem olduğundan dolayı optimizasyon algoritmalarının ikili versiyonları kullanılmıştır.

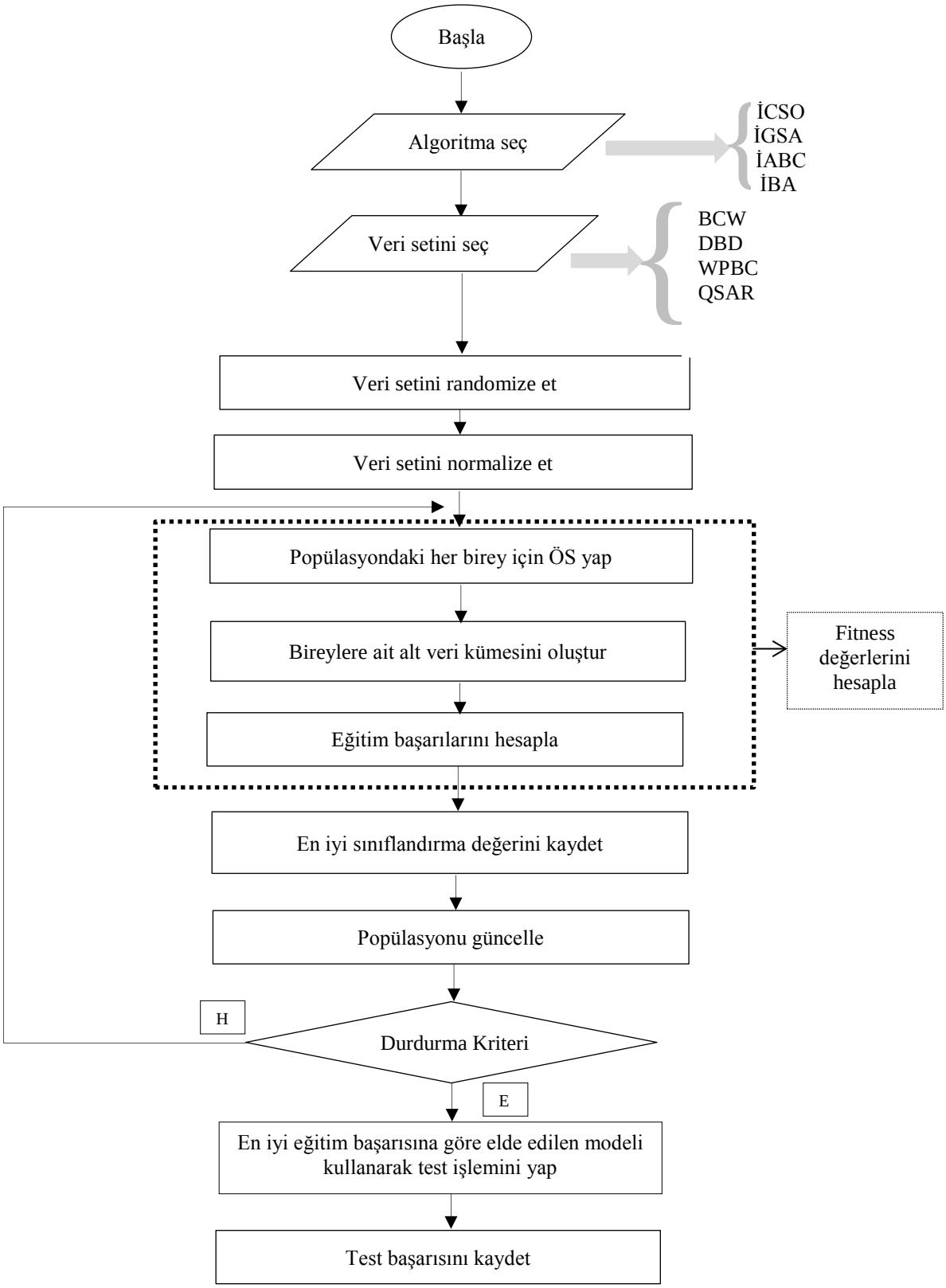
Önerilen sistemde üzerinde çalışılan veri setinde kaç tane özellik olduğu dikkate alınarak ikili bir sistem oluşturulmuştur. Kullanılan bireylerin boyutları veri kümelerindeki özellik sayısına eşittir. Popülasyondaki her bir bireyin ilgili boyutları 0 ve 1 değerlerini almış şekilde ikili (binary) bir sistem oluşturulmuştur. Örnek olarak

"wisconsin göğüs kanseri" veri kümesine ilişkin başlangıçta oluşturulan rastgele bir birey aşağıda gösterilmiştir. Örnek birey 9 boyutlu bir yapıya sahiptir. Bu popülasyonda, bireylerin "0" değerleri o özelliğin seçilmediğini, "1" değerleri de ilgili veri setinde kendisine karşılık gelen özelliklerin seçildiğini göstermektedir.

Birey = [0 1 1 1 0 1 0 1 0];

Başlangıçta verilen popülasyon sayısına göre popülasyon kümesi 0 ve 1'ler rastsal olacak şekilde oluşturulur. Oluşturulan popülasyon kümesindeki her bir birey DVM'ye gönderilir. DVM birimi içerisinde *c* ve *γ* parametrelerine göre eğitim başarısı test edilir. Her bir iterasyonda eğitim başarısının artırılması hedef alınarak grid arama yöntemiyle en uygun c ve γ parametreleri bulunur. Popülasyon içerisinde en yüksek eğitim başarısına sahip bireye doğru diğer bireylerin yaklaşması sağlanır. Algoritma önceden belirlenmiş olan iterasyon sayısına kadar işlem bu şekilde devam ettirilir. En son olarak sisteme ait en yüksek eğitim başarısı, eğitilmiş modelin test başarısı, en yüksek eğitim başarısının yakalandığı c ve γ değerleri, her bir çalışma (run-time) için geçen toplam işlemci süresi (CPU time), duyarlılık (sensitivity), özgüllük (specifity), artı yorum gücü (positive predictive value), eksi yorum gücü (negative predictive value), bireyin ikili sistem karşılığı ve ÖS'deki özellik sayısı kaydedilir. Bu değerlere ilişkin ortalama ve standart değerler de ayrıca kayıt altına alınır. Bunlara ilaveten sistemin her bir iterasyondaki eğitim başarısının gözlemlenmesi amacıyla iteratif olarak algoritmadan elde edilen popülasyondaki en iyi bireye ait eğitim başarısı ve ikili sistem karşılığı da kaydedilir. Önerilen ÖS sisteminin akış diyagramı Çizelge 5.1'de verilmiştir.

Çizelge 5.1. Önerilen ÖS sisteminin akış diyagramı



6. VERİ AYRIKLAŞTIRMA TEKNİKLERİ VE ÖNERİLEN YÖNTEM

6.1. Veri Ayrıklaştırma İşlemi

Ayrıklaştırma veya İngilizce adıyla discretization (veya binning) sayısal verilerin kategorik karşılıklarına dönüştürülmesi işlemidir. Örneğin yaş değişkeninin değerlerini 12-19, 20-39, 40-59 ve 60-79 gibi aralıklara bölünerek gruplama işlemi buna bir örnek olarak verilebilir. Sayısal değişkenler genellikle sıklık dağılımlarına göre ayrıklaştırılır. Bu işlem sayesinde verideki gürültü ve doğrusalsızlık (non-linearity) azaltılarak tahminleme modelinin kesinliliği arttırılabilir. Bunun yanında aykırı gözlemlerin tespiti, geçersiz veya eksik numerik değerlerin tespiti de kolaylaştırılır.

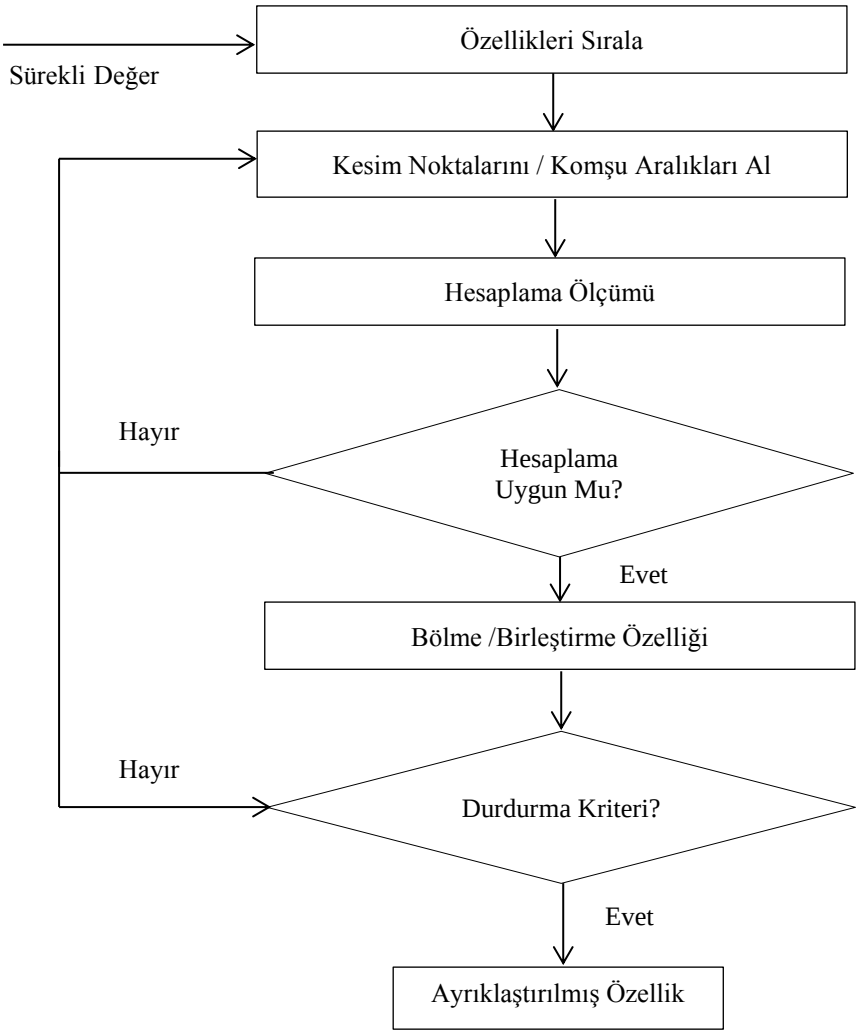
Naive Bayes ağları, tümevarım kuralları, birliktelik kuralları ve bazı veri madenciliği algoritmaları ayrık değerlere ihtiyaç duymaktadır (Grzymala-Busse ve Stefanowski 2001). Bundan dolayı makine öğrenme işleminden önce, her bir sürekli özelliğin bir dizi aralık değerleriyle (interval) oluşturulan ayrık bir özellik uzayı içinde haritalanması gerekmektedir. Ayrıklaştırma olarak bilinen bu işlem sadece bazı öğrenme metotlarının sürekli değerler yerine ayrık değerlere ihtiyaç duyması gibi bir sebepten dolayı veri ön işlemede önemli ve gerekli bir süreçtir. Aynı zamanda bir dizi intervaller kullanarak yeniden oluşturulmuş verinin yorumlanmasının daha kolay olması, azaltılmış seviyedeki veri ile hesaplama işleminin daha hızlı gerçekleşmesi ve birbirleriyle alakası olmayan ilişkileri bulma gibi önemli sebeplerden dolayı ayrıklaştırma önemli bir süreç olarak kabul edilir (Liu ve ark 2002).

6.2. Karakteristik Bir Ayrıklaştırma İşlemi

Ayrıklaştırma tek değişkenli ya da çok değişkenli olabilir. Tek değişkenli ayrıklaştırma her hangi bir zamanda tek bir sürekli değişkeni dikkate alırken çok değişkenli ayrıklaştırma da eş zamanlı bir biçimde çoklu özelliklere göre işlem yapar. Karakteristik ayrıklaştırma literatürde tek değişkenli ayrıklaştırma olarak bilinir.

Tipik bir ayrıklaştırma işlemi Çizelge 6.1'de görüldüğü gibi genel olarak dört adımdan oluşur: (1) *Sıralama*, ayrıklaştırılacak özelliğe ilişkin sürekli değerlerin sıralanması, (2) *Hesaplama*, birleştirmek için komşu aralıkların hesaplanması veya bölmek için bölme noktaların hesaplanması, (3) bazı kriterlere göre sürekli değerlerin aralıklarını *bölme* ya da *birleştirme*, (4) *Durdurma kriteri*, belli bir kritere göre işlemin sonlanmasını ifade etmektedir.

Çizelge 6.1. Ayrıklaştırma İşlemi



6.2.1. Sıralama

Bir özellik için sürekli değerler artan ya da azalan şekilde sıralıdır. Ayrıklaştırma ile birlikte sıralama işlemine uygulamada dikkat edilmezse sıralama hesaplama açısından çok maliyetli olabilir. Uygun bir sıralama algoritmasını seçerek ayrıklaştırma işlemini hızlandırmak önemlidir.

6.2.1. Kesim noktası seçme

Sıralama işleminden sonra, ayrıklaştırma işlemindeki sonraki adım bir dizi sürekli değerleri ayıracak en iyi kesim noktalarını bulmak veya birleştirmek için en iyi komşu aralıkları bulmaktır. Tipik bir hesaplama fonksiyonu sınıf etiketiyle birlikte birleştirme ve bölme korelasyonunu belirler.

6.2.3. Bölme (split) ve birleştirme (merge)

Bölme işlemi için kesim noktalarının hesaplanması ve en iyi kesim noktasının seçilmesi gerekmektedir. Bu işlemden sonra, sürekli değer dizisinin iki kısma bölünmesi gerekir. Bu işlem sonlanma kriteri sağlanana kadar, ayrıklaştırma artan şekilde yeniden oluşan her bölme ile devam eder. Benzer şekilde birleştirme işleminde, her bir adımda komşu aralıklar bir araya getirilerek en iyi çift şeklindeki aralıklar bulunur. Sonlanma kriteri sağlanıncaya kadar ayrıklaştırma azalan sayıdaki aralıklarla devam eder.

6.2.4. Durdurma kriteri

Durdurma kriteri başlangıçta sabit bir aralık sayısı gibi basit bir değer olabileceği gibi aynı zamanda bir fonksiyonu hesaplama gibi daha karışık bir değer de olabilir.

6.3. Veri Ayrıklaştırma Yöntemlerinin Sınıflandırılması

Veri ayrıklaştırma yöntemleri özelliklerine göre belirli gruplara göre listelenmektedir. Dougherty ve ark (1995) bu gruplandırmayı denetimli ve denetimsiz, dinamik ve statik, yerel ve genel olmak üzere üç grupta ifade ederken, Yang ve Webb (2005), veri ayrıklaştırma yöntemlerini Şekil 6.1'deki gibi dokuz grupta sınıflandırmıştır.

1. Denetimli ve Denetimsiz: Kesim noktalarını elde etmek için sınıf bilgisi kullanmayanlar denetimsiz, sınıf bilgisini kullanan yöntemler ise denetimli yöntemler olarak ifade edilmektedir. Eşit genişlikli, EFD yöntemleri sınıf bilgisine ihtiyaç duymadan ayrıklaştırma yaptığından denetimsiz; ID3, 1RD gibi yöntemler ise sınıf bilgisini kullandıkları için denetimli yöntemler olarak ifade edilmektedir.

2. Parametrik ve Parametrik olmayan: Her ayrıklaştırma işlemi için kullanıcı tarafından bir parametre girilmesini isteyen yöntemler parametrik, parametre gerektirmeyen yöntemler ise parametrik olmayan yöntemler olarak ifade edilir.

3. Hiyerarşik ve Hiyerarşik olmayan: Hiyerarşik yöntemlerde kesim noktası bulunurken artırımlı bir süreç izlenerek hiyerarşik bir yapı oluşturur. Bölümlemeli ve birleştirmeli bir işlem izleyen yöntemler bu grup içinde yer almaktadır. ChiMerge, Chi2 bu hiyerarşik yöntemlere örnek verilebilir.

4. Tekli ve Çoklu değişkenli: Çok değişkenli yöntemler ayrıklaştırma sürecinde özellikler arası ilişkiyi göz önünde bulundururken, tek değişkenli yöntemler herhangi bir nitelik değerine atıf yapmaksızın işlemini gerçekleştirir.

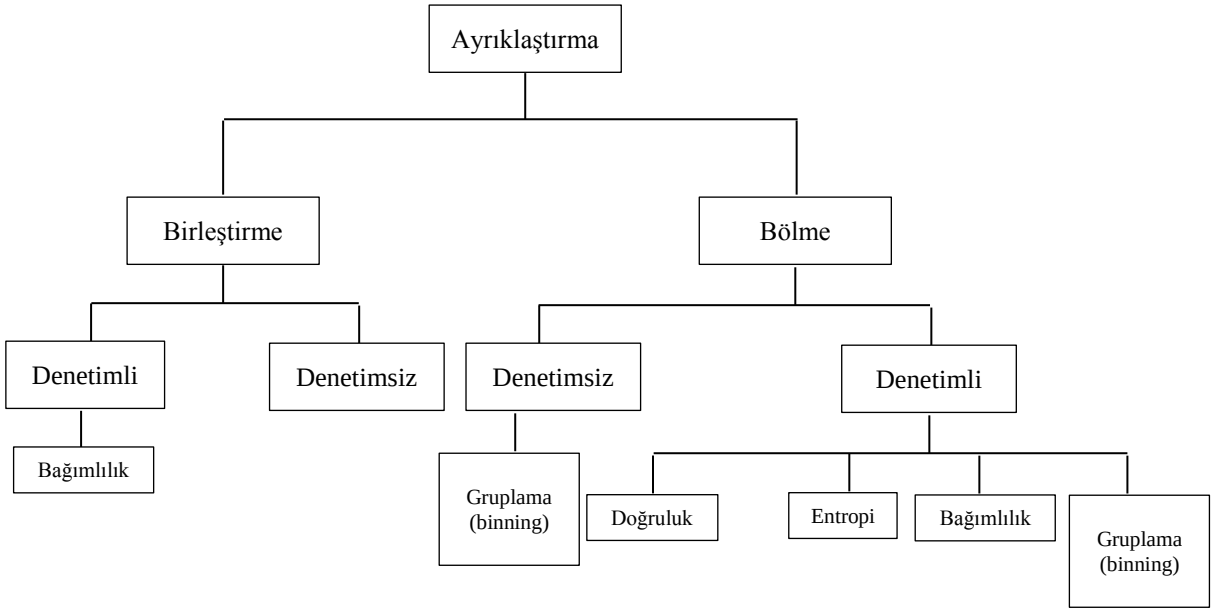
5. Ayrık ve Ayrık olmayan: Ayrık metotlar değer aralıklarını birbiri ile çakışmayacak şekilde ayrık aralıklar oluşturur.

6. Genel ve Yerel: Genel ayrıklaştırma yöntemleri sadece bir sınıflandırma işlemi için eşleştirme fonksiyonu oluştururlar. Yerel yöntemlerde ise durum tam bunun tersidir. Yerel yöntemler örnek uzayın lokal bölümlerine uygulanarak bölümler oluşturulurken, genel yöntemler ise her bir özelliğin diğer özelliklerden bağımsız bir şekilde farklı bir gruba dahil olacağı n boyutlu örnek uzay üzerinde bir ağ meydana getirirler (Dougherty ve ark 1995).

7. Gayretli ve Tembel: Gayretli yöntemler sınıflandırma işi için öncelikli bir eşleme fonksiyonu oluştururken, tembel yöntemler sadece ihtiyaç anında bu işi gerçekleştirir. Buna ek olarak tembel yöntemler tüm sürekli nitelikler için bir test veri kümesi verilene kadar işlem yapmadan beklerken her bir özelliğin çevresinde sadece iki kesim noktası oluşturur (Hsu ve ark 2003).

8. Sıralı ve Sembolik: Sıralı yöntemler nicelden nitel sıralı niteliklere bir eşleme fonksiyonu oluştururlar.

9. Bulanık ve Bulanık olmayan: Bulanık yöntemler bulanık eşleme oluştururken, değerler farklı genişliklerdeki birden fazla aralık içinde yer alabilirler. Bunun dışında bulanık ayrıklaştırma, sürekli nitelik değerindeki küçük değişimlerin bu niteliğe ait olasılıklar üzerindeki etkisinin küçük olması fikrine dayanmaktadır.



Şekil 6.1. Ayrıklaştırma yöntemleri için hiyerarşik bir çatı

6.4. Veri Ayrıklaştırmada Kullanılan Bazı Yöntemler

Literatürde oldukça sık kullanılan veri ayrıklaştırma yöntemlerinden bazıları aşağıda açıklanmıştır.

6.4.1. Eşit genişlikli ayrıklaştırma

Eşit Genişlikli Ayrıklaştırma yönteminde gözlemlenen herhangi bir değer için maksimum ve minimum değerler bulunur (Kurgan ve Cios 2004). Bu değerler, eşit genişlikte olan ve kullanıcı tarafından belirlenmiş *n* adet aralığa yerleştirilir (Dougherty ve ark 1995). Herhangi bir sınıf bilgisi kullanmamasından dolayı denetimsiz ayrıklaştırma yöntemleri arasında yer almaktadır. EWD yöntemi bu tez çalışmasında kullanılmıştır.

Bu yöntemin sade olması diğer yöntemler arasında kendisini daha cazip kılsa da (Boulle 2005), aralık genişliklerinin belirlenme işlemi ve ardından eşit aralıklara bölme işlemi doğru yapılamadığı durumda bu yöntem ayrıklaştırma işlemi sonrasında büyük bir veri kaybına sebep olabilir (Wong ve Chiu 1987).

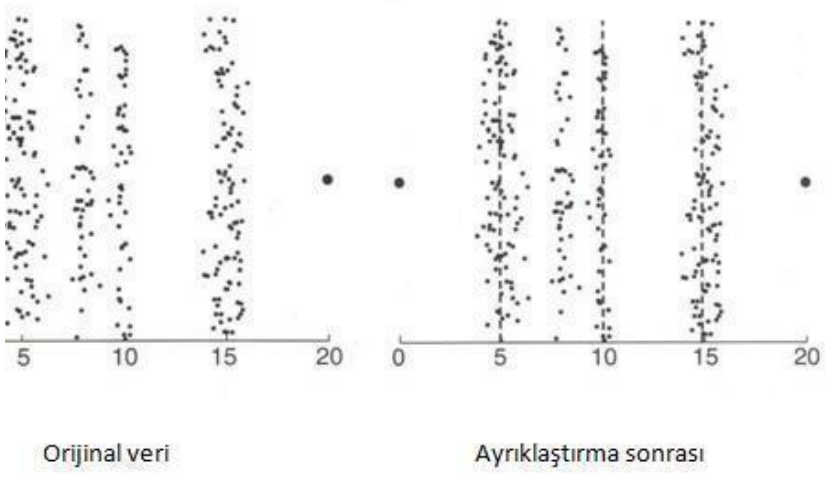
Yang ve Webb (2002), çalışmalarında bu yöntemi ve oluşacak aralıkların kesim noktası olarak ifade edilen sınır noktalarını şu şekilde tanımlamışlardır:

$$w = (vmaks - vmin) / t \tag{6.1}$$

$$vmin + w, vmin + 2w, \ldots, vmin + (t-1)w \tag{6.2}$$

Bu formülleri kısaca açıklamak gerekirse, denklem 6.1'de maksimum (vmaks) ve minimun (vmin) değerler arasındaki farkın belirlenen aralık sayısına (t) bölünmesi ile elde edilen w değişkeni aralıkların genişliklerini ifade etmektedir. Denklem 6.2'deki minumum değerden başlamak üzere genişlik değerinin ve katlarının eklenmesi ile elde edilen değerler de aralıkların bitiminde yer alan ve kesim noktası olarak ifade edilen değerlerin büyüklüklerini göstermektedir.

Şekil 6.2'de örnek olarak orijinal veri ile EWD yönteminin uygulanması sonucu elde edilen verilerin görünümü yer almaktadır.



Şekil 6.2. EWD yöntemi sonrası verilerin görünümü (Tan ve ark 2006)

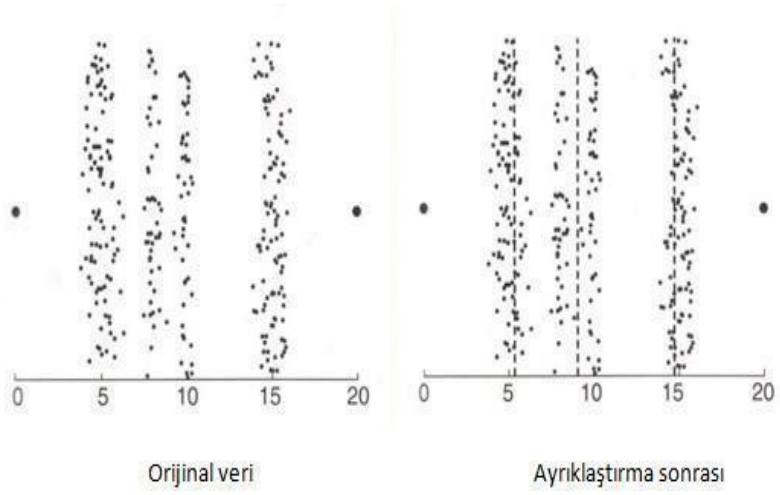
6.4.2. Eşit frekanslı ayrıklaştırma

Eşit frekanslı ayrıklaştırma yöntemi de EWD yöntemi gibi herhangi bir sınıf bilgisi kullanmadan uygulanan denetimsiz ayrıklaştırma yöntemlerinden birisidir (Biba ve ark 2007). Bu yöntemde ilk olarak veri setinde yer alan özellikler artan bir şekilde dizilmektedir. Bu işlemin ardından bu nitelik değerleri daha önceden belirlenmiş olan her bir aralığa eşit sayıda gelecek şekilde aralıklara yerleştirilmektedir(Mehta ve ark 2005, Kotsiantis ve Kanellopoulos 2006). EFD yöntemi bu tez çalışmasında kullanılmıştır.

Dougherty ve ark (1995), çalışmalarında bu yöntemin işleyişini biraz daha formülize ederek n adet aralığa, her aralıkta x/n adet olmak üzere eşit sayıda nitelik değeri gelecek şekilde yerleştirilme işlemi şeklinde yöntemi belirlemişlerdir. Burada x,

sürekli değişkenin sahip olduğu örnek sayısını ifade etmektedir. Bu işlem n-1 aralığın oluşmasını sağlarken EWD'de de olduğu gibi, her aralığa ait sınır değerler, bir başka ifadeyle kesim noktaları meydana gelmektedir. Komşu kesim noktaları arası alan normalize edilmiş histogramda 1/n değerine olabildiğince yakın olacaktır (Khan ve Revett 2004).

Şekil 6.3'de orijinal veri ile EFD yönteminin uygulanması sonucu elde edilen verilerin görünümü yer almaktadır.



Şekil 6.3. EFD sonrası verilerin görünümü (Tan ve ark 2006)

6.5. Ayrıklaştırma İçin Önerilen Sistem

Bu çalışmada literatürde en çok rastlanan ayrıklaştırma yöntemlerinden olan EWD ve EFD yöntemleri kullanılmıştır. Her iki ayrıklaştırma yöntemi için de aynı sistem önerilmiştir. En iyi ayrıklaştırılmış veri kümesini oluşturmak/bulmak için dört farklı optimizasyon algoritması kullanılmıştır. Bu algoritmalar yukarıda belirtildiği gibi *ABC, CSO, BA ve GSA* algoritmalarıdır. Bu algoritmaları ayrıklaştırma işlemini gerçekleştirmek için veri setleri üzerine uygulamadan önce ayrıklaştırma için gerekli olmayan özellikler veri kümesinden çıkartılmıştır. Bununla birlikte ayrıklaştırma işlemi sadece giriş değerlerine göre yapıldığından veri kümesindeki sınıf bilgisi de veri kümesinden çıkartılmış ve veri setleri ayrıklaştırmaya hazır hale getirilmiştir. Veri ön işlemeden sonra her bir veri kümesinin özellik sayısına göre başlangıç popülasyonundaki bireylerin boyutu belirlenmiştir. Bireylere ait her bir boyuta 2 ila 10

değerleri de dahil olmak üzere bu sayılar arasında rastgele bir sürekli değer seçilerek popülasyon oluşturulmuştur. Burada minimum 2 seçilerek ilgili özelliğin en az 2 aralığa bölüneceği en fazla 10 seçilerek de maksimum 10 aralıkla özelliğin ifade edileceği söylenmiştir. Ayrıklaştırma işlemi sürekli değerler yerine ayrık değerlere göre çalıştığından dolayı ayrıklaştırma aşamasında bireylere ait sürekli değerler *round* fonksiyonu ile yuvarlanarak ayrık bir uzay oluşturulmuş olur.

Önerilen sistemde üzerinde çalışılan veri setinde kaç tane özellik olduğu dikkate alınarak sürekli değerlere sahip bir sistem oluşturulmuştur. Kullanılan bireylerin boyutları veri kümelerindeki özellik sayısına eşittir. Popülasyondaki her bir bireyin ilgili boyutları 2 ila 10 değerleri arasındaki sürekli bir değer almış şekilde sürekli değerlere sahip bir sistem oluşturulmuştur. Örnek olarak *"wisconsin göğüs kanseri"* veri kümesine ilişkin başlangıçta oluşturulan rastgele bir birey aşağıda gösterilmiştir. Örnek birey 9 boyutlu bir yapıya sahiptir. Bu popülasyonda, her hangi bir bireye ait sürekli değerler ve onların round edilmiş versiyonu aşağıda sunulmuştur.

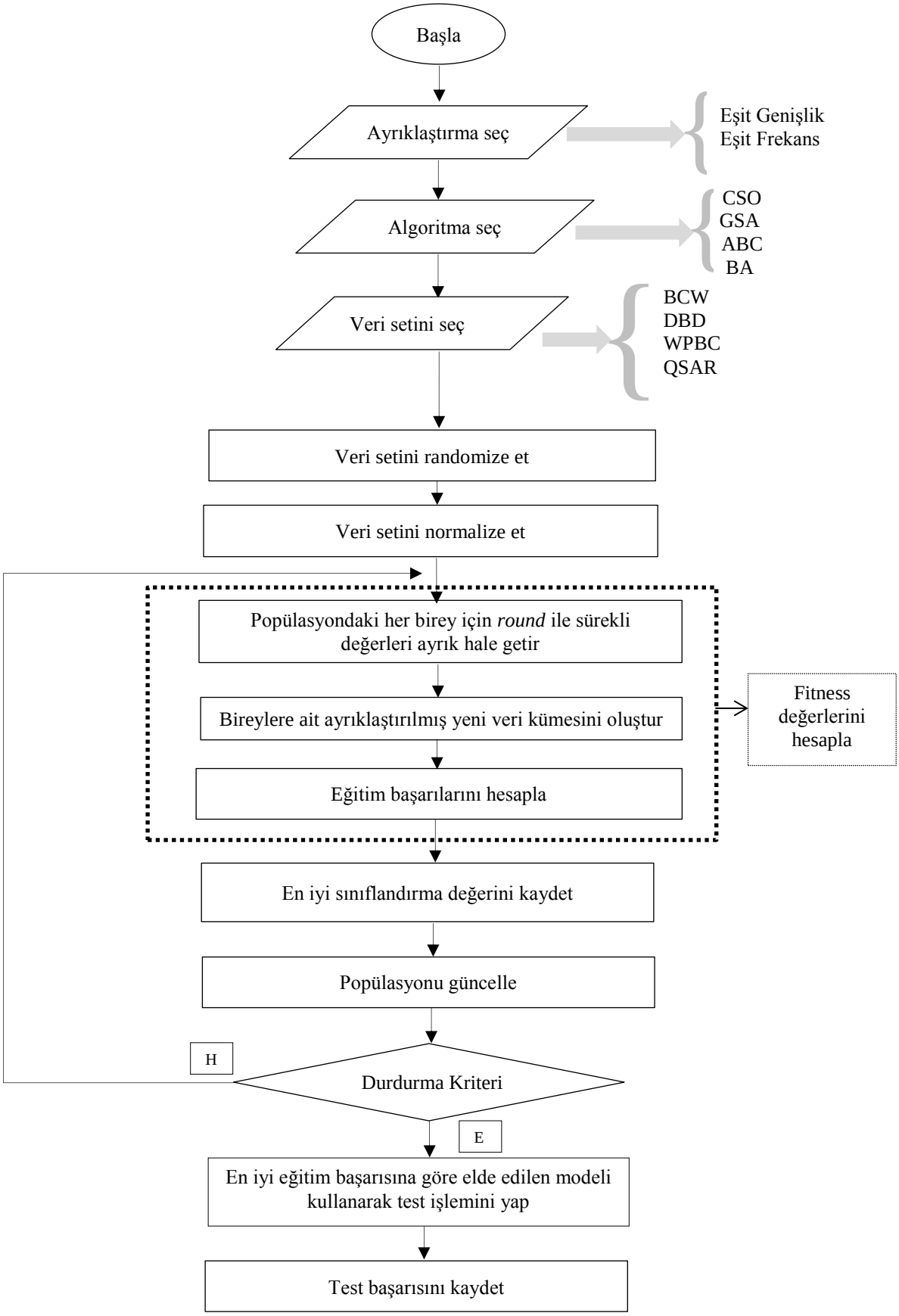
6,057	9,989	6,296	6,197	2,236	5,252	3,927	8,916	3,104

6	10	6	6	2	5	4	9	3

Başlangıçta verilen popülasyon sayısına göre popülasyon kümesi 2 ila 10 değerleri arasında sürekli değerler şeklinde rastsal olarak oluşturulur. Oluşturulan popülasyon kümesindeki her bir birey *round* fonksiyonu ile veriler ayrık hale getirilerek DVM'ye gönderilir. DVM birimi içerisinde *c* ve γ parametrelerine göre eğitim başarısı test edilir. Her bir iterasyonda eğitim başarısının artırılması hedef alınarak grid arama yöntemiyle en uygun c ve γ parametreleri bulunur. Popülasyon içerisinde en yüksek eğitim başarısına sahip bireye doğru diğer bireylerin yaklaşması sağlanır. Algoritma önceden belirlenmiş olan iterasyon sayısına kadar işlem bu şekilde devam ettirilir. En son olarak sisteme ait en yüksek eğitim başarısı, eğitilmiş modelin test başarısı, en yüksek eğitim başarısının yakalandığı c ve γ değerleri, her bir çalışma (run-time) için geçen toplam işlemci süresi (CPU time), duyarlılık (duyarlılık), özgüllük (specifity), artı yorum gücü (predictive value), eksi yorum gücü (negative predictive), bireyin sürekli değerlerden oluşan vektör karşılığı kaydedilir. Bu değerlere ilişkin ortamala ve standart değerler de ayrıca kayıt altına alınır. Bunlara ilaveten sistemin her bir iterasyondaki eğitim başarısının gözlemlenmesi amacıyla iterasyonel olarak algoritmadan elde edilen

popülasyondaki en iyi bireye ait eğitim başarısı ve sürekli değerlerden oluşan vektör karşılığı da kaydedilir. Önerilen sistemin akış diyagramı Çizelge 6.2'de gösterilmiştir.

Çizelge 6.2. Önerilen Ayrıklaştırma sisteminin akış diyagramı



7. TEST VE DEĞERLENDİRME

7.1. Deneysel Çalışmalar

Özellik seçiminde ve ayrıklaştırma işlemlerinde veri kümesi olarak UCI veri tabanından *BCW, QSAR, DRD* ve *WPBC* isimli farklı sayıda kayda ve özelliğe sahip veri kümeleri seçilmiştir. Veri setlerine ait özellik sayısı ve kayıt sayısı Çizelge 7.1'de verilmiştir. Burada verilen özellik sayısı veri ön işleme yapılmadan önceki durumdaki özellik sayısını belirtmektedir. Veri ön işleme aşamasında ise *BCW* veri setinde ilk özellik *Id* değeri olup gereksiz bir özellik olduğundan veri kümesinden çıkartılmıştır. Bu veri setinde 16 eksik veri olup eksik verilerin tümü 6. özellikte yer almaktadır. Bu eksik verilerin tamamlanabilmesi için de ilgili özelliğin ortalaması alınarak eksik hücreler tamamlanmıştır. Ortalama değer 3.5 olarak hesaplanmış ve ilgili veri setinde tüm değerler tam sayı değerleri olduğu için sayı bir üst tam sayıya yuvarlanarak elde edilen 4 değeri ile eksik hücreler doldurulmuştur. *QSAR* ve *DRD* veri kümelerinde ise eksik değer bulunmamaktadır. Son olarak; *WPBC* veri setinde ilk özellik *Id* değeri olup gereksiz bir özellik olduğundan veri kümesinden çıkartılmıştır. Bu veri setinde 4 eksik veri olup eksik verilerin tümü *lymph* isimli özellikte yer almaktadır. Bu eksik verilerin tamamlanabilmesi için de ilgili özelliğin ortalaması alınarak eksik hücreler tamamlanmıştır. Ortalama değer 3.2 olarak hesaplanmış ve ilgili veri setinde tüm değerler tam sayı değerleri olduğu için sayı bir alt tam sayıya yuvarlanarak elde edilen 3 değeri ile eksik kayıtlar doldurulmuştur.

İlk olarak her bir veri kümesi kendi içerisinde her bir çalışma anında (run-time) eğitim ve test işlemlerinin gerçek başarısını görmek için (randomize edilmiştir) karıştırılmıştır. Tüm veri setleri [0-1] aralığına normalize edilmiştir. Verilerin %80'i eğitim verisi, %20'i de test verisi olarak değerlendirilmiştir. Elde edilen sonuçlar kayıt altına alınmıştır.

Çizelge 7.1. Veri setlerine ilişkin genel özellikler

Veri Seti	Özellik Sayısı	Kayıt Sayısı	Eksik Özellik
Breast-Cancer-Wisconsin	10	699	16 adet
QSAR Biodegradation	42	1055	Yok
Diabetic Retinopathy Debrecen	20	1151	Yok
Breast Cancer Wisconsin (Prognostic) (WPBC)	34	198	4 adet

Önerilen sistemlerin gerçekleştirilmesi için "svmlib" ismindeki destek vektör makinası kütüphanesi kullanılmıştır (Matlab için geliştirilmiş Destek Vektör Makinesi Tools). Önerilen sistemlerde en iyi çözüm kaydedilirken aynı zamanda DVM'nin parametreleri olan c ve γ değerleri ve ilgili çözümde hangi özelliklerin seçildiği de sonuç olarak kaydedilmektedir. c ve γ değerlerinin her ikisi için 2^x şeklinde üstel olarak test edilmiştir. Burada x ifadesine -5 ila +5 arasında tam sayı değerleri verilmiştir. x değerleri -5'den +5'e kadar birer artırılmıştır. Dolayısıyla önerilen sistemde c ve γ değerleri 2^{-5} ila 2^{+5} arasında değerler almıştır.

CSO algoritmasında parametre değerleri; habitat sayısı 20, beta değeri 3 ve hayatta kalma oranı 0.25 olarak kullanılmıştır. GSA'nın parametre değerleri; kütle sayısı 20, G_0 1 olarak kullanılmıştır. BA'da A,r,r_0 0.5 olarak, F_{min} 0, $F_{max,}$ 1, alfa ve lamda parametreleri de 0.9 olarak alınmıştır.

Her üç algoritma da farklı iterasyonlarda 5 kez çalıştırılmıştır (robustness). Bu iterasyonlar sırasıyla 100, 200 ve 300 olarak belirlenmiştir.

Yukarıda bahsedildiği gibi her bir algoritmadaki başlangıç popülasyon değerleri sırasıyla 10,20,30,40 ve 50 olarak test edilmiştir.

7.2. Tanı Testlerinde Kullanılan Ölçütler

Tanı testleri göreli olarak değerlendirilirler ve testin özellikleri ölçüt yöntemlerinin sonuçlarına bağlı hesaplamalarla elde edilir. Özellikler hasta/sağlam ve test (+)/test (-) birlikte ele alınarak değerlendirilir. Tanı ve tarama testlerinin değerlendirilmesinde genellikle kullanılan ölçütler duyarlılık, özgüllük, artı yorum gücü ve eksi yorum gücü değerleridir. Bu ölçütler için Çizelge 7.2'den yararlanılarak testlerin sınır değerleri belirlenmektedir. Önerilen sistemin akış diyagramı Çizelge 7.2'de gösterilmiştir (Çamlica ve Disci 2008).

Çizelge 7.2. Duyarlılık, özgüllük, artı yorum gücü ve eksi yorum gücü değerlerinin hesaplanması

	Referans Test		
Yeni Test	**Hasta**	**Sağlam**	**Toplam**
+	Gerçek Pozitif (GP)	Yalancı pozitif (YP)	GP+YP
-	Yalancı negatif (YN)	Gerçek negatif (GN)	YN+GN
Toplam	GP+YN	YP+GN	GP+YP+ YN+GN

Duyarlılık, gerçekte sağlık problemine sahip olan bireyler içinde tarama testinin hasta bireyleri bulabilme özelliği olarak tanımlanmaktadır. Duyarlılık kısaca bir tarama testinin hasta bulma yeteneği olarak da tanımlanabilir.

Duyarlılık = (GP / (GP+YN)) * 100 = % olarak ifade edilir.

Özgüllük, gerçekte sağlıklı olan bireyler içinde testin sağlıklı bireyleri bulabilme özelliği olarak tanımlanmaktadır. Özgüllük kısaca bir tarama testinin sağlıklı bireyleri bulma yeteneği olarak da tanımlanabilir.

Özgüllük = (GN / (YP+GN)) * 100 = % olarak ifade edilir.

Pozitif yorum gücü, sınıflandırma sonucu sağlık problemli olarak bulunan bireyler içinde gerçekten sağlık problemine sahip olanların yüzdesidir.

Pozitif Yorum Gücü = (GP / (GP+YP)) * 100 = % olarak ifade edilir.

Negatif yorum gücü, sınıflandırma sonucu sağlıklı olarak bulunan bireyler içinde gerçekten sağlıklı olanların yüzdesidir.

Negatif Yorum Gücü = (GN / (YN+GP)) * 100 = % olarak ifade edilir.

7.3. Özellik Seçimi İçin Test Sonuçları

7.3.1. WPBC veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar

WPBC veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.3'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.1 ile Şekil 1.6 arasında grafiksel olarak gösterilmiştir. Çizelge 7.3 ile 7.26 arasındaki tablolarda ifade edilen *P.S.* popülasyon sayısını belirtirken, *İ.S.* ise iterasyon sayısını göstermektedir.

Çizelge 7.3. ÖS kullanılarak WPBC veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	71,50	76,50	**77,00**	70,00	70,50	74,50	76,00	**790**	66,00	**75,50**	68,50	75,00
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	68,00	**80,50**	75,50	77,50	67,50	72,00	**78,50**	70,50	70,50	77,00	**82,50**	75,50
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	70,50	75,50	**76,00**	72,50	75,00	75,00	**80**,00	69,50	73,50	79,50	**83,00**	74,50
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	68,00	74,00	**76,00**	70,50	71,50	76,50	72,00	**80,00**	69,50	76,50	71,50	**77,00**
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	66,00	**78,00**	76,00	75,50	70,50	**77**,00	73,00	72,50	**77,00**	75,50	73,00	74,00

WPBC veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.4'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.7 ile Şekil 1.18 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.4. ÖS kullanılarak WPBC veri kümesi üzerinde algoritmalardan elde edilen istatiksel sonuçlar

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	**9,30**	0,00	0,00	0,00	**16,10**	0,00	1,67	4,00	**19,90**	2,80	0,00	0,00
	Özgüllük	86,40	**100**	100	99,30	88,10	**100**	100	100	79,30	98,80	**100**	100
	Artı Yorum Gücü	**19,20**	0,00	0,00	0,00	**25,90**	0,00	0,00	0,00	**21,20**	0,00	0,00	0,00
	Eksi Yorum Gücü	**79,70**	76,50	77,00	70,40	77,50	74,50	75,90	**78,90**	**78,20**	76,30	68,50	75,00
20	**Duyarlılık**	**35,80**	0,00	0,00	0,00	**19,70**	0,00	0,00	0,00	**16,70**	0,00	0,00	1,82
	Özgüllük	80,10	**100**	100	100	80,00	**100**	100	100	88,30	**100**	100	100
	Artı Yorum Gücü	**37,10**	0,00	0,00	0,00	**22,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	79,20	**80,50**	75,50	77,50	**79,30**	72,00	78,50	70,50	77,00	77,00	**82,50**	75,40
30	**Duyarlılık**	**18,20**	0,00	0,00	2,86	**20,60**	0,00	0,00	1,25	**33,80**	6,67	0,00	3,33
	Özgüllük	89,20	**100**	100	100	90,30	**100**	100	98,50	80,50	**99,40**	99,40	99,40
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**36,70**	0,00	0,00	0,00	**21,80**	0,00	0,00	0,00
	Eksi Yorum Gücü	75,50	75,50	**76,00**	72,40	**81,10**	75,00	80,00	70,00	**87,30**	79,70	83,40	74,60
40	**Duyarlılık**	**22,60**	0,00	3,48	0,00	**17,00**	0,00	0,00	2,50	**45,60**	0,00	0,00	2,50
	Özgüllük	83,80	**100**	100	100	88,50	**100**	100	98,80	78,10	**100**	100	99,40
	Artı Yorum Gücü	**39**,00	0,00	0,00	0,00	**29,60**	0,00	0,00	0,00	**40,20**	0,00	0,00	0,00
	Eksi Yorum Gücü	75,60	74,00	**75,70**	70,50	77,10	76,50	72,00	**80,70**	**80,90**	76,50	71,50	77,30
50	**Duyarlılık**	**21,70**	0,00	2,86	2,220	**22,00**	1,43	0,00	0,00	**38,30**	2,86	0,00	0,00
	Özgüllük	77,70	**100**	99,40	98,70	82,10	**100**	100	98,00	85,10	99,30	99,40	**100**
	Artı Yorum Gücü	**22,70**	0,00	0,00	0,00	**18,20**	0,00	0,00	0,00	**28,90**	0,00	0,00	0,00
	Eksi Yorum Gücü	**79,40**	78,00	76,30	76,20	**81,80**	76,80	73,00	73,60	**87,60**	75,80	73,40	74,00

7.3.2. BCW veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar

BCW veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.5'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.19 ile Şekil 1.24 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.5. ÖS kullanılarak BCW veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	94,60	92,90	94,00	**95,40**	95,40	93,90	**95,60**	95,30	94,30	94,40	**94,90**	94,00
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**95,90**	94,00	95,10	95,30	**95,30**	94,90	93,60	93,70	**95,10**	94,30	94,40	94,10
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	93,70	93,00	**96,40**	95,70	**95,30**	95,00	94,70	92,40	**95,30**	94,60	93,00	95,30
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	94,00	95,70	94,90	**96,70**	**95,60**	94,30	95,30	94,60	94,40	94,60	95,30	**95,70**
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	95,30	**96,60**	95,00	95,30	95,70	94,10	95,70	**96,40**	**94,90**	94,00	93,60	94,40

BCW veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.6'da gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.25 ile Şekil 1.36 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.6. ÖS kullanılarak BCW veri kümesi üzerinde algoritmalardan elde edilen istatiksel sonuçlar

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	96,50	91,00	**99,50**	96,90	97,00	97,10	97,70	**98,30**	94,60	95,10	**98,80**	94,50
	Özgüllük	93,50	93,60	91,30	**94,70**	**94,70**	92,20	94,30	93,90	94,10	**94,20**	92,90	93,70
	Artı Yorum Gücü	89,80	89,20	84,90	**91,20**	90,10	86,60	**91,30**	88,40	88,80	89,30	87,50	**89,60**
	Eksi Yorum Gücü	97,90	95,10	**99,80**	98,10	98,50	98,40	98,60	**99,10**	97,50	97,20	**99,30**	96,80
20	**Duyarlılık**	95,80	93,70	**98,20**	95,40	97,80	94,00	**98,40**	94,20	97,20	92,40	**99,30**	99,20
	Özgüllük	**95,90**	94,10	93,50	95,30	93,90	**95,30**	90,80	93,60	94,10	**95,40**	91,60	91,70
	Artı Yorum Gücü	**92,30**	90,50	88,90	91,20	89,30	**90,70**	85,80	88,10	89,90	**91,70**	87,50	85,40
	Eksi Yorum Gücü	97,80	96,40	**99,10**	97,60	98,90	96,90	**99,10**	97,00	98,40	95,80	**99,50**	99,50
30	**Duyarlılık**	94,00	91,30	**99,20**	98,60	96,70	93,20	**99,30**	95,70	97,20	94,80	**99,60**	96,20
	Özgüllük	93,50	93,80	**94,80**	94,30	94,40	**95,80**	92,70	90,30	94,00	94,40	89,90	**95,00**
	Artı Yorum Gücü	88,30	87,80	**91,70**	89,20	89,60	**92,30**	86,90	84,80	**91,00**	89,60	82,70	89,90
	Eksi Yorum Gücü	97,00	95,60	**99,50**	99,40	98,50	96,50	**99,50**	98,20	98,40	97,40	**99,80**	98,00
40	**Duyarlılık**	94,10	97,20	**99,60**	98,50	95,70	95,20	**98,40**	97,70	96,10	94,10	**99,60**	96,80
	Özgüllük	93,80	94,80	92,30	**95,90**	**95,50**	93,80	93,70	92,80	93,40	94,80	92,90	**95,10**
	Artı Yorum Gücü	88,30	91,70	87,50	**92,50**	**92,50**	89,60	89,40	88,70	89,80	90,30	88,40	**91,70**
	Eksi Yorum Gücü	97,20	98,40	**99,80**	99,10	97,60	97,30	**99,10**	98,60	97,60	96,90	**99,70**	98,20
50	**Duyarlılık**	98,40	95,20	**98,50**	97,30	96,00	93,80	**100**	99,50	94,50	92,60	**98,70**	96,10
	Özgüllük	93,80	**97,20**	93,00	94,10	**95,40**	94,20	93,60	94,90	**95,10**	94,80	91,00	93,60
	Artı Yorum Gücü	88,30	**94,50**	89,20	90,60	**91,60**	89,00	88,40	90,70	**91,50**	91,30	84,70	89,00
	Eksi Yorum Gücü	**99,00**	97,70	99,00	98,40	98,10	96,90	**100**	99,80	96,80	95,70	**99,30**	97,70

7.3.3. QSAR veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar

QSAR veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.7'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.37 ile Şekil 1.42 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.7. ÖS kullanılarak QSAR veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	99,20	99,90	**100**	99,80	99,20	99,90	**100**	99,80	99,10	99,90	**100**	99,90
	Test Başarısı	74,20	75,20	**77,70**	74,40	77,50	**78,10**	72,50	76,20	75,70	**77,00**	76,30	75,90
20	**Eğitim Başarısı**	**100**	99,90	100	99,90	**100**	99,90	100	99,90	**100**	99,90	100	100
	Test Başarısı	75,60	**78,10**	78,00	76,30	75,60	77,00	75,50	**77,10**	74,00	**78,50**	75,50	78,40
30	**Eğitim Başarısı**	99,60	**100**	100	99,90	99,50	**100**	100	99,90	99,70	**100**	100	99,90
	Test Başarısı	**77,40**	77,20	75,60	75,20	77,30	**78,00**	75,60	76,20	76,00	**80,80**	74,00	75,50
40	**Eğitim Başarısı**	99,60	**100**	100	99,90	99,60	**100**	100	99,90	99,70	99,90	**100**	99,90
	Test Başarısı	73,10	77,40	**77,70**	75,20	75,70	75,00	75,60	**76,20**	**76,80**	76,10	76,30	75,50
50	**Eğitim Başarısı**	99,70	99,90	**100**	99,90	99,70	99,90	**100**	99,90	99,70	99,90	**100**	99,90
	Test Başarısı	74,50	**78,10**	75,60	78,10	74,90	**77,00**	75,60	77,00	76,70	**78,50**	75,50	78,50

QSAR veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.8'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.43 ile Şekil 1.54 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.8. ÖS kullanılarak QSAR veri kümesi üzerinde algoritmalardan elde edilen istatiksel sonuçlar

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	49,00	**62,00**	50,10	52,50	53,50	**64,70**	34,60	51,90	51,10	**60,30**	47,40	50,70
	Özgüllük	87,30	81,80	**92,20**	85,30	89,70	84,90	**92,50**	87,70	88,00	85,00	**91,30**	89,10
	Artı Yorum Gücü	66,80	63,00	**76,80**	64,50	**72,40**	68,50	70,40	66,00	67,70	66,10	**74,10**	70,60
	Eksi Yorum Gücü	76,80	**81,20**	78,20	78,40	79,30	**82,80**	73,10	79,70	78,50	**81,50**	70,00	77,80
20	**Duyarlılık**	40,00	**66,60**	44,20	51,60	44,50	**62,40**	38,70	55,90	40,40	**63,00**	42,70	51,40
	Özgüllük	93,50	83,50	**94,40**	88,20	90,60	85,00	**94,50**	88,90	91,50	86,60	**92,20**	91,70
	Artı Yorum Gücü	75,70	65,50	**79,50**	68,00	68,90	69,70	**77,90**	740	70,50	70,80	73,40	**74,80**
	Eksi Yorum Gücü	75,50	**84,20**	77,80	79,10	77,50	**80,50**	75,10	78,30	74,80	**81,80**	76,10	79,40
30	**Duyarlılık**	51,10	**64,50**	40,00	46,20	46,60	**62,70**	44,50	49,90	49,80	**70,60**	40,40	48,30
	Özgüllük	90,40	83,70	**93,50**	90,00	**91,00**	85,50	90,60	89,20	89,30	85,60	**91,50**	90,00
	Artı Yorum Gücü	72,00	67,50	**75,70**	71,10	**70,50**	67,90	68,90	69,70	69,80	69,90	70,50	**72,80**
	Eksi Yorum Gücü	79,10	**82,40**	75,50	76,50	79,10	**82,40**	77,50	78,40	78,00	**85,90**	74,80	76,80
40	**Duyarlılık**	41,40	**65,10**	50,10	46,20	43,10	**63,10**	44,50	49,90	47,10	**62,80**	47,40	48,30
	Özgüllük	90,10	84,30	**92,20**	90,00	**91,30**	81,20	90,60	89,20	**92,00**	82,60	91,30	90,00
	Artı Yorum Gücü	69,40	70,40	**76,80**	71,10	**70,40**	64,00	68,90	69,70	**74,80**	64,40	74,10	72,80
	Eksi Yorum Gücü	74,10	**81,00**	78,20	76,50	77,00	**80,80**	77,50	78,40	77,30	**81,70**	77,00	76,80
50	**Duyarlılık**	45,10	**66,60**	40,00	66,60	47,10	**62,40**	44,50	62,40	48,20	**63,00**	42,70	63,00
	Özgüllük	90,70	83,50	**93,50**	83,50	88,40	85,00	**90,60**	85,00	91,20	86,60	**92,20**	86,60
	Artı Yorum Gücü	73,20	65,50	**75,70**	65,50	66,40	**69,70**	68,90	69,70	**73,60**	70,80	73,40	70,80
	Eksi Yorum Gücü	75,00	**84,20**	75,50	84,20	77,50	**80,50**	77,50	80,50	77,40	**81,80**	76,10	81,80

7.3.4. DRD veri kümesi kullanılarak ÖS'ye ait deneysel sonuçlar

DRD veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.9'da gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.55 ile Şekil 1.60 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.9. ÖS kullanılarak DRD veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	97,50	99,50	**99,80**	99,00	97,80	99,60	**99,80**	99,40	98,70	99,60	**99,90**	99,50
	Test Başarısı	66,80	**67,50**	63,50	63,20	65,80	**66,80**	60,30	59,20	63,10	**63,50**	62,70	61,00
20	**Eğitim Başarısı**	97,30	99,70	**99,80**	99,40	98,90	99,70	**99,80**	99,60	99,30	99,70	**99,80**	99,50
	Test Başarısı	65,10	63,60	64,00	**68,70**	**65,70**	61,00	62,70	63,60	62,80	63,20	62,60	**69,40**
30	**Eğitim Başarısı**	98,60	99,60	**99,90**	99,60	98,90	99,70	**99,80**	99,60	99,20	**99,80**	99,80	99,70
	Test Başarısı	**64,20**	63,20	62,20	61,10	**62,70**	61,00	61,70	60,20	63,50	59,70	**64,80**	60,90
40	**Eğitim Başarısı**	99,00	99,70	**99,80**	99,70	98,90	**99,80**	99,00	99,80	99,60	**99,90**	98,80	99,70
	Test Başarısı	58,00	60,00	**61,60**	61,60	62,00	61,80	**65,60**	64,40	62,50	56,90	**68,60**	61,60
50	**Eğitim Başarısı**	98,60	99,70	**99,80**	99,50	99,40	**99,80**	99,80	99,70	99,30	**99,80**	99,80	99,70
	Test Başarısı	62,30	61,50	60,90	**66,80**	61,00	58,60	**64,20**	62,60	62,30	58,60	**64,20**	62,80

DRD veri kümesi kullanılarak ÖS yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.10'da gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-1'de Şekil 1.61 ile Şekil 1.72 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.10. ÖS kullanılarak DRD veri kümesi üzerinde algoritmalardan elde edilen istatiksel sonuçlar

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	65,50	61,70	**69,60**	65,70	65,20	61,8	**67,10**	63,60	63,50	62,00	**67,70**	58,80
	Özgüllük	68,40	**74,70**	56,60	60,40	66,40	**73,00**	53,00	54,50	62,50	**65,00**	57,00	63,90
	Artı Yorum Gücü	70,80	**72,90**	64,60	64,50	67,80	**73,60**	61,00	60,20	66,40	65,80	64,20	**67,40**
	Eksi Yorum Gücü	63,10	**63,80**	62,00	61,70	**64,10**	61,40	59,50	58,00	59,60	**61,20**	60,70	55,30
20	**Duyarlılık**	70,80	61,40	69,60	**70,90**	69,30	57,50	**69,90**	61,80	67,60	62,10	66,90	**70,00**
	Özgüllük	59,00	66,00	56,60	**66,50**	61,60	64,90	54,20	**65,50**	57,60	63,40	57,20	**68,90**
	Artı Yorum Gücü	65,50	68,50	67,40	**70,50**	**68,10**	65,00	63,50	64,50	64,30	66,70	65,70	**72,00**
	Eksi Yorum Gücü	64,80	59,40	59,00	**67,30**	**63,00**	57,80	61,10	62,80	61,60	59,10	58,60	**66,70**
30	**Duyarlılık**	68,90	56,10	**69,10**	60,50	65,70	**67,40**	67,00	64,70	65,40	58,50	**70,00**	62,20
	Özgüllük	58,50	**70,40**	54,00	62,20	**59,60**	53,20	55,90	55,30	61,10	**62,20**	58,80	59,20
	Artı Yorum Gücü	65,90	**67,30**	63,40	64,20	**63,50**	63,40	62,50	60,70	**67,30**	65,20	66,20	63,20
	Eksi Yorum Gücü	**61,80**	60,10	60,30	58,60	**61,90**	57,00	60,70	59,50	59,00	55,40	**63,00**	58,10
40	**Duyarlılık**	65,80	50,30	**67,60**	66,30	67,30	68,50	68,60	**69,00**	64,40	57,10	**69,60**	64,60
	Özgüllük	49,80	**71,60**	54,70	56,60	56,50	53,20	52,70	**58,80**	**60,20**	57,70	50,70	58,30
	Artı Yorum Gücü	58,10	**66,90**	63,00	63,30	61,90	63,50	63,10	**66,80**	**64,80**	63,80	61,40	64,40
	Eksi Yorum Gücü	57,90	55,80	59,60	**60,10**	**62,20**	0,00	58,90	61,20	59,90	**62,90**	54,40	58,50
50	**Duyarlılık**	64,60	55,60	**65,30**	65,00	67,50	65,60	**68,00**	66,10	**69,10**	65,60	68,00	64,30
	Özgüllük	59,70	**69,10**	55,80	68,80	53,10	51,40	**59,90**	58,30	54,50	51,40	59,90	**60,70**
	Artı Yorum Gücü	64,70	69,70	63,10	**70,50**	63,40	64,70	**66,60**	65,40	63,80	64,70	**66,60**	63,80
	Eksi Yorum Gücü	59,60	55,00	58,20	**63,50**	57,60	**62,30**	61,30	59,10	60,20	**62,30**	61,30	61,20

7.4. Eşit Genişlikli Ayrıklaştırma Yöntemi İçin Test Sonuçları

7.4.1. WPBC verisi kullanılarak EWD yöntemine ait deneysel sonuçlar

WPBC veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.11'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.1 ile Şekil 2.6 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.11. EWD kullanılarak WPBC veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	I.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	78,50	72,50	**79,50**	74,50	77,00	75,50	73,00	**81,50**	75,50	**76,50**	74,00	72,50
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	76,50	**77,00**	76,00	73,50	76,00	**77,50**	71,00	73,00	75,50	77,00	**81,50**	77,50
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**76,00**	75,00	76,00	75,50	**80,00**	79,00	78,00	78,50	76,50	**78,00**	76,50	76,00
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**80,50**	77,00	79,00	78,50	74,50	**78,50**	77,00	73,00	**80,50**	73,00	74,50	75,00
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**82,50**	75,50	82,00	77,50	**78,50**	74,50	74,50	75,50	76,00	**79,00**	76,00	75,50

WPBC veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.12'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.7 ile Şekil 2.18 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.12. EWD kullanılarak WPBC veri kümesi üzerinde algoritmaların istatiksel sonuçları

	I.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	**12,20**	0,00	0,00	0,00	**4,44**	0,00	0,00	0,00	**9,19**	0,00	0,00	0,00
	Özgüllük	98,80	**100**	100	100	**100**	100	100	100	97,30	**100**	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0**,00	0,00	0,00	0,00	**0**,00	0,00	0,00	0,00
	Eksi Yorum Gücü	78,90	72,50	**79,50**	74,50	76,80	75,50	73,00	**81,50**	76,40	**76,50**	74,00	72,50
20	**Duyarlılık**	**7,22**	0,00	0,00	0,00	**10,00**	0,00	0,00	0,00	**5,64**	0,00	0,00	0,00
	Özgüllük	99,40	**100**	100	100	98,60	**100**	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	76,60	**77,00**	76,00	73,50	76,20	**77,50**	71,00	73,00	75,10	77,00	**81,50**	77,50
30	**Duyarlılık**	**6,52**	0,00	0,00	0,00	**6,94**	0,00	0,00	0,00	**11,60**	0,00	0,00	0,00
	Özgüllük	99,40	**100**	100	100	98,00	**100**	99,40	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0**,00	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**76,10**	75,00	76,00	75,50	**80,90**	79,00	78,40	78,50	75,80	**78,00**	76,50	76,00
40	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**4,76**	0,00	0,00	0,00	**12,70**	0,00	0,00	0,00
	Özgüllük	99,40	**100**	100	100	**100**	100	100	100	98,20	**100**	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**80,90**	77,00	79,00	78,50	74,30	**78,50**	77,00	73,00	**81,50**	73,00	74,50	75,00
50	**Duyarlılık**	**15,70**	0,00	0,00	0,00	**5,56**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	99,40	99,40	**100**	100	98,10	**100**	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**82,60**	75,90	82,00	77,50	**79,50**	74,50	74,50	75,50	76,00	**79,00**	76,00	75,50

7.4.2. BCW veri kümesi kullanılarak EWD yöntemine ait deneysel sonuçlar

BCW veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.13'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.19 ile Şekil 2.24 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.13. EWD kullanılarak BCW veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	99,90	100	100	**100**	100	100	100
	Test Başarısı	94,40	92,10	92,90	**95,30**	**91,40**	88,90	83,90	87,30	91,60	**94,00**	90,70	87,70
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	93,60	90,40	**94,30**	83,40	93,90	89,40	**95,40**	93,10	**91,10**	88,40	87,90	88,90
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**94,40**	93,40	89,90	88,40	93,70	**93,90**	79,90	86,00	93,00	92,70	92,70	**94,10**
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	94,00	93,10	92,90	**96,00**	92,40	91,10	90,60	**94,30**	94,00	93,70	88,60	**95,10**
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	94,10	**95,00**	90,70	82,60	93,90	**94,90**	92,30	86,90	93,90	87,60	**95,00**	87,30

BCW veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.14'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.25 ile Şekil 2.36 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.14. EWD kullanılarak BCW veri kümesi üzerinde algoritmaların istatiksel sonuçları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	99,60	99,20	**100**	100	**99,30**	80,00	60,00	78,70	**99,60**	99,10	99,60	78,80
	Özgüllük	91,80	88,10	88,80	**92,80**	87,30	94,00	**95,90**	91,90	87,40	91,70	85,70	**92,70**
	Artı Yorum Gücü	**86,10**	82,10	83,70	0,00	80,50	**89,80**	0,00	0,00	80,40	0,00	80,60	**87,90**
	Eksi Yorum Gücü	99,80	99,50	**100**	100	**99,40**	92,70	86,60	92,30	**99,80**	99,60	99,70	92,10
20	**Duyarlılık**	**100**	79,20	100	60,50	**100**	81,50	100	99,60	**100**	79,50	100	80,40
	Özgüllük	90,40	94,10	91,30	**94,80**	90,80	92,70	**93,10**	89,80	86,80	**93,00**	80,20	93,00
	Artı Yorum Gücü	83,70	**89,70**	86,50	0,00	84,40	**88,60**	0,00	0,00	79,00	0,00	77,70	**88,50**
	Eksi Yorum Gücü	**100**	94,00	100	87,00	**100**	93,90	100	99,80	**100**	93,20	100	93,30
30	**Duyarlılık**	99,20	99,60	**100**	80,90	**99,60**	99,50	60,00	80,00	99,20	**100**	99,60	99,20
	Özgüllük	91,90	89,50	84,80	**92,30**	90,50	90,80	89,90	**90,90**	89,60	88,50	89,50	**91,40**
	Artı Yorum Gücü	**87,00**	85,90	79,10	0,00	85,00	**85,20**	0,00	0,00	84,10	0,00	82,40	**86,10**
	Eksi Yorum Gücü	99,50	99,70	**100**	93,60	**99,80**	99,80	86,60	92,00	99,50	**100**	99,80	99,50
40	**Duyarlılık**	**100**	98,90	99,10	100	99,30	**100**	99,10	99,00	**100**	99,60	99,60	99,60
	Özgüllük	90,90	89,80	90,00	**93,80**	88,30	87,00	86,30	**91,60**	90,70	91,00	82,40	**92,70**
	Artı Yorum Gücü	85,00	85,10	83,50	**89,80**	83,50	78,80	80,00	**86,50**	85,50	83,70	76,80	**88,10**
	Eksi Yorum Gücü	**100**	99,30	99,60	100	99,50	**100**	99,50	99,60	**100**	99,80	99,70	99,80
50	**Duyarlılık**	**100**	98,10	98,80	61,70	**100**	99,60	98,70	80,40	**99,60**	79,60	99,60	80,70
	Özgüllük	91,00	93,40	87,20	**93,60**	90,30	**92,40**	89,00	90,50	90,70	91,70	**92,80**	92,50
	Artı Yorum Gücü	85,40	**89,10**	78,70	88,70	85,70	**87,50**	81,80	85,60	85,50	86,90	87,00	**88,70**
	Eksi Yorum Gücü	**100**	98,90	99,30	86,80	**100**	99,80	99,30	92,90	**99,80**	92,80	99,80	92,00

7.4.3. QSAR veri kümesi kullanılarak EWD yöntemine ait deneysel sonuçlar

QSAR veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.15'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.37 ile Şekil 2.42 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.15. EWD kullanılarak QSAR veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	96,30	99,50	97,30	**100**	97,00	99,70	97,70	**100**	97,30	99,80	97,30
	Test Başarısı	71,00	69,40	68,90	**71,10**	**73,30**	70,70	68,60	69,70	**74,00**	66,80	70,60	68,80
20	**Eğitim Başarısı**	**100**	97,60	98,60	97,70	**100**	97,50	99,00	98,20	**100**	97,20	98,80	97,70
	Test Başarısı	**74,90**	70,00	67,80	67,80	**74,90**	70,00	68,20	71,40	**74,60**	71,40	68,10	70,00
30	**Eğitim Başarısı**	**100**	97,70	99,20	98,20	**100**	97,60	99,10	98,10	**100**	97,60	99,10	98,20
	Test Başarısı	**74,40**	67,90	68,00	70,40	**72,80**	69,40	69,50	69,50	**70,80**	70,10	66,50	68,60
40	**Eğitim Başarısı**	**100**	97,90	99,50	98,30	**100**	97,60	99,50	98,10	**100**	97,70	99,50	97,90
	Test Başarısı	**73,60**	70,00	68,20	71,40	**74,00**	68,40	70,20	69,20	**73,70**	69,40	66,60	68,40
50	**Eğitim Başarısı**	**100**	98,20	99,50	98,20	**100**	98,00	99,60	98,50	**100**	97,80	99,60	98,10
	Test Başarısı	**77,40**	70,00	67,90	68,10	**74,30**	68,20	67,30	69,10	**71,80**	69,60	68,80	70,10

QSAR veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.16'da gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.43 ile Şekil 2.54 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.16. EWD kullanılarak DRD veri kümesi üzerinde algoritmaların istatiksel sonuçları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	**42,30**	15,80	7,71	12,80	**51,50**	23,20	6,40	11,90	**42,90**	14,50	11,60	15,40
	Özgüllük	87,40	96,90	**98,60**	97,90	84,30	96,40	**99,30**	99,20	87,50	95,50	**97,90**	96,70
	Artı Yorum Gücü	64,30	71,50	**82,30**	75,20	62,60	75,90	78,00	**87,20**	60,40	66,40	**72,70**	71,40
	Eksi Yorum Gücü	**73,30**	69,20	68,70	70,90	**77,70**	69,90	68,30	68,80	**78,00**	67,10	70,60	68,70
20	**Duyarlılık**	**42,00**	15,90	12,00	14,10	**48,80**	13,40	11,20	12,70	**47,70**	20,50	11,10	16,10
	Özgüllük	89,70	97,40	**97,50**	96,40	89,40	96,90	**98,80**	97,80	87,30	96,70	**98,10**	97,10
	Artı Yorum Gücü	64,10	**75,00**	69,80	70,80	71,40	68,00	**82,80**	72,20	63,80	**75,40**	72,70	74,00
	Eksi Yorum Gücü	**77,70**	69,60	67,50	67,90	**76,10**	70,10	67,40	71,40	**78,00**	71,10	67,70	69,80
30	**Duyarlılık**	**40,90**	14,20	10,80	15,20	**45,50**	16,70	14,50	12,00	**35,90**	13,30	10,20	12,90
	Özgüllük	91,10	96,50	**98,30**	97,60	87,30	95,50	97,30	**97,40**	90,10	**98,40**	97,50	97,70
	Artı Yorum Gücü	68,90	68,90	74,10	**74,20**	66,10	63,90	**71,90**	68,40	65,90	**79,70**	69,20	74,60
	Eksi Yorum Gücü	**75,80**	67,90	67,60	70,10	**74,90**	69,90	69,20	69,60	**72,20**	69,60	66,40	68,20
40	**Duyarlılık**	**44,70**	12,70	9,34	17,50	**47,80**	13,80	10,20	12,90	**44,40**	15,90	6,95	14,40
	Özgüllük	88,00	97,30	**98,00**	98,00	88,30	98,20	**98,50**	98,10	87,80	97,20	**98,60**	98,20
	Artı Yorum Gücü	64,90	70,00	71,30	**81,60**	68,60	**81,20**	76,70	79,00	64,10	73,70	72,80	**81,70**
	Eksi Yorum Gücü	**76,00**	70,10	68,20	70,60	**75,70**	67,60	70,00	68,70	**76,70**	69,00	66,50	67,50
50	**Duyarlılık**	**50,30**	15,00	5,990	13,10	**42,30**	9,40	8,240	12,30	**37,80**	14,70	10,50	14,50
	Özgüllük	89,60	**98,10**	97,50	97,50	90,10	98,20	**98,40**	97,40	87,70	98,10	**98,80**	96,60
	Artı Yorum Gücü	68,60	**81,00**	52,10	75,30	68,50	**74,60**	72,60	71,10	61,70	80,60	**84,60**	66,60
	Eksi Yorum Gücü	**80,10**	69,40	68,50	67,60	**76,00**	67,90	67,10	69,00	**75,00**	68,80	68,20	70,40

7.4.4. DRD veri kümesi kullanılarak EWD yöntemine ait deneysel sonuçlar

DRD veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.17'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.55 ile Şekil 2.60 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.17. EWD kullanılarak DRD veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	89,50	96,90	90,50	**100**	89,90	97,50	91,80	**100**	91,90	97,70	91,80
	Test Başarısı	**62,80**	57,00	58,40	60,30	**63,60**	58,30	56,50	55,80	**63,70**	57,30	57,60	57,10
20	**Eğitim Başarısı**	**100**	92,50	94,00	93,20	**100**	92,10	94,10	92,10	**100**	91,70	94,50	91,70
	Test Başarısı	**64,20**	59,80	58,60	56,90	**64,40**	56,60	57,90	59,10	**63,30**	59,50	56,80	59,50
30	**Eğitim Başarısı**	**100**	91,90	95,50	93,20	**100**	92,70	95,90	91,90	**100**	91,90	95,70	92,80
	Test Başarısı	**62,60**	56,10	56,50	56,80	**63,40**	56,90	57,10	57,00	**63,80**	61,00	58,10	58,90
40	**Eğitim Başarısı**	**100**	92,60	96,00	92,50	**100**	92,70	96,50	92,00	**100**	91,10	96,30	92,80
	Test Başarısı	**64,40**	57,60	58,30	60,30	**61,50**	61,40	57,70	59,00	**62,60**	59,80	59,00	58,00
50	**Eğitim Başarısı**	**100**	92,30	96,80	92,90	**100**	92,70	96,90	92,80	**100**	92,60	97,00	92,90
	Test Başarısı	**65,00**	59,80	56,20	59,60	**64,30**	58,10	56,00	56,80	**64,00**	58,20	55,60	55,80

DRD veri kümesi kullanılarak EWD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.18'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-2'de Şekil 2.61 ile Şekil 2.72 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.18. EWD kullanılarak DRD veri kümesi üzerinde algoritmaların istatiksel sonuçları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	64,90	77,40	**87,00**	82,00	67,10	80,30	**87,70**	79,10	62,10	83,20	**89,70**	79,20
	Özgüllük	**60,40**	33,70	25,20	34,70	**60,00**	33,10	21,40	29,60	**65,50**	28,10	22,00	32,40
	Artı Yorum Gücü	**64,30**	57,50	57,60	59,50	**63,90**	57,90	55,70	55,80	**67,50**	56,60	56,10	57,20
	Eksi Yorum Gücü	61,10	55,90	62,00	**62,40**	**63,40**	59,20	60,20	56,00	60,00	60,50	**66,20**	57,70
20	**Duyarlılık**	68,90	82,50	**84,10**	82,30	64,20	81,90	**84,00**	83,30	64,50	**86,10**	83,60	79,80
	Özgüllük	**59,20**	33,30	28,10	29,40	**64,60**	27,70	28,70	32,50	**61,80**	31,80	27,10	35,00
	Artı Yorum Gücü	**64,60**	59,60	58,50	55,70	**67,10**	56,50	56,90	57,60	**64,80**	56,90	56,10	59,90
	Eksi Yorum Gücü	**64,10**	62,00	59,80	60,70	61,80	57,30	62,40	**63,80**	61,50	**68,70**	58,60	58,60
30	**Duyarlılık**	61,80	81,20	81,10	**83,40**	64,90	81,00	**85,80**	80,10	65,50	85,30	**85,90**	82,40
	Özgüllük	**63,80**	28,50	28,40	28,80	**61,50**	30,20	25,30	32,30	**62,50**	30,50	27,80	32,90
	Artı Yorum Gücü	**66,30**	55,60	56,50	55,20	**66,20**	56,20	56,50	55,80	**65,80**	60,70	56,80	57,70
	Eksi Yorum Gücü	59,30	58,30	57,20	**62,70**	60,20	58,80	**62,10**	60,50	62,70	62,40	**64,10**	62,50
40	**Duyarlılık**	64,20	83,20	**88,70**	83,40	61,80	83,90	**85,50**	81,30	62,30	76,40	**85,40**	82,40
	Özgüllük	**64,80**	30,80	25,70	33,10	**61,30**	37,10	28,10	33,00	**62,90**	39,60	28,20	31,90
	Artı Yorum Gücü	**65,90**	55,80	56,30	59,70	**64,20**	59,70	56,20	58,80	**65,80**	60,80	58,40	57,00
	Eksi Yorum Gücü	63,10	63,10	**67,70**	62,80	58,80	**67,60**	64,10	60,30	59,50	57,50	**62,90**	61,30
50	**Duyarlılık**	64,50	84,10	**87,90**	84,80	65,70	84,20	**86,70**	80,50	67,60	82,20	**86,90**	80,40
	Özgüllük	**66,00**	32,00	21,30	30,20	**62,90**	29,50	20,80	31,50	**59,90**	31,00	22,60	28,80
	Artı Yorum Gücü	**69,40**	58,60	55,10	59,20	**67,70**	57,20	55,80	55,80	**64,50**	57,80	54,30	55,50
	Eksi Yorum Gücü	60,80	**63,80**	61,40	63,20	60,70	**62,80**	58,10	60,50	**63,20**	59,90	61,90	56,90

7.5. Eşit Frekanslı Ayrıklaştırma Yöntemi İçin Test Sonuçları

7.5.1. WPBC verisi kullanılarak EFD yöntemine ait deneysel sonuçlar

WPBC veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.19'da gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.1 ile Şekil 3.6 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.19. EFD kullanılarak WPBC veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	77,00	72,50	**79,50**	74,50	76,00	75,50	73,00	**81,50**	75,00	**76,50**	74,00	72,50
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	75,50	75,50	**76,50**	73,50	74,50	**78,00**	74,50	77,50	74,00	**77,50**	71,50	77,50
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	75,00	76,50	**79,00**	75,50	80,00	81,50	**83,00**	78,00	73,50	**78,00**	71,00	76,00
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**81,00**	76,50	75,50	72,50	73,50	76,50	**80,50**	75,50	**80,00**	73,50	76,00	74,50
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**80,50**	77,50	75,00	74,00	**79,00**	75,50	78,50	78,00	76,00	**79,00**	74,00	77,50

WPBC veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.20'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.7 ile Şekil 3.18 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.20. EFD kullanılarak WPBC veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	77,00	72,50	**79,50**	74,50	76,00	75,50	73,00	**81,50**	75,00	**76,50**	74,00	72,50
20	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	75,50	75,50	**76,50**	73,50	74,50	**78,00**	74,50	77,50	74,00	**77,50**	71,50	77,50
30	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	75,00	76,50	**79,00**	75,50	80,00	81,50	**83,00**	78,00	73,50	**78,00**	71,00	76,00
40	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**81,00**	76,50	75,50	72,50	73,50	76,50	**80,50**	75,50	**80,00**	73,50	76,00	74,50
50	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**80,50**	77,50	75,00	74,00	**79,00**	75,50	78,50	78,00	76,00	**79,00**	74,00	77,50

7.5.2. BCW verisi kullanılarak EFD yöntemine ait deneysel sonuçlar

BCW veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.21'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.19 ile Şekil 3.24 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.21. EFD kullanılarak BCW veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	64,00	**66,90**	63,70	66,00	**67,30**	66,90	66,00	64,30	63,70	**64,60**	63,40	63,10
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	65,00	66,30	**67,90**	66,00	63,30	65,90	**68,40**	63,70	**66,90**	66,10	65,60	65,60
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	62,60	63,30	**69,00**	64,30	65,60	65,00	**67,90**	65,90	65,30	64,10	66,30	**66,40**
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**68,40**	64,40	66,40	65,90	**66,90**	65,60	65,10	64,70	65,40	**65,70**	62,60	65,10
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	64,10	63,30	65,70	**67,70**	**67,00**	64,10	66,60	65,90	65,70	**67,60**	65,90	65,00

BCW veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.22'de gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.25 ile Şekil 3.36 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.22. EFD kullanılarak BCW veri kümesi üzerinde algoritmaların istatiksel sonuçları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	0,00	0,00	**0,37**	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	64,00	**66,90**	63,70	66,00	**67,30**	66,90	66,00	64,30	63,70	**64,60**	63,40	63,10
20	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	65,00	66,30	**67,90**	66,00	63,30	65,90	**68,40**	63,70	**66,90**	66,10	65,6	65,60
30	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	62,60	63,30	**69,00**	64,30	65,60	65,00	**67,90**	65,90	65,30	64,10	66,30	**66,40**
40	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**68,40**	64,40	66,40	65,90	**66,90**	65,60	65,10	64,70	65,40	**65,70**	62,60	65,10
50	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,44**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	64,10	63,30	65,70	**67,70**	**67,00**	64,10	66,60	65,90	65,70	**67,60**	65,90	65,00

7.5.3. QSAR veri kümesi kullanılarak EFD yöntemine ait deneysel sonuçlar

QSAR veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.23'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.37 ile Şekil 3.42 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.23. EFD kullanılarak QSAR veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	64,10	66,10	66,40	**66,90**	66,60	64,80	66,40	**69,00**	**69,90**	64,60	66,40	68,60
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	66,70	65,70	**68,00**	67,30	**67,30**	64,80	65,10	63,80	66,10	65,30	**67,30**	66,60
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	**69,40**	66,00	65,70	65,10	65,00	**69,60**	66,60	64,70	63,40	68,00	65,50	**68,30**
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	69,40	**70,00**	67,30	68,20	65,00	64,80	**69,30**	68,40	63,40	66,40	**67,80**	66,40
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	66,50	64,40	**69,70**	66,00	66,60	**67,20**	65,40	66,70	66,10	66,70	65,40	**67,40**

QSAR veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.24'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.43 ile Şekil 3.54 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.24. EFD kullanılarak DRD veri kümesi üzerinde algoritmalardan elde edilen istatiksel sonuçlar

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Özgüllük	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	64,10	66,10	66,40	**66,90**	66,60	64,80	66,40	**69,00**	**69,90**	64,60	66,40	68,60
20	**Duyarlılık**	0,54	0,00	**10,20**	0,00	**1,17**	0,00	0,31	0,00	0,30	0,00	**6,00**	0,00
	Özgüllük	**100**	100	98,80	100	99,70	**100**	100	100	**100**	100	99,70	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	66,70	65,70	**67,40**	67,30	**67,30**	64,80	65,10	63,80	66,00	65,30	**66,90**	66,60
30	**Duyarlılık**	**5,72**	2,25	0,84	3,23	2,34	**7,32**	0,83	1,34	3,42	4,42	3,10	**10,30**
	Özgüllük	99,90	**100**	100	99,60	**100**	100	100	100	99,80	99,60	**99,90**	98,90
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**69,00**	65,70	65,60	64,90	64,70	**68,80**	66,50	64,60	63,00	67,70	65,20	**68,00**
40	**Duyarlılık**	5,72	**9,19**	2,93	4,35	2,34	0,50	**4,84**	4,37	3,42	2,22	**7,51**	2,79
	Özgüllük	99,90	99,30	99,70	**100**	**100**	100	99,70	99,90	99,80	99,90	99,90	**100**
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	69,00	**69,70**	67,10	67,80	64,70	64,80	**69,00**	68,10	63,00	66,30	**66,90**	66,20
50	**Duyarlılık**	0,00	0,00	**7,48**	0,00	0,00	0,00	**8,63**	0,00	0,00	0,00	**8,63**	0,00
	Özgüllük	**100**	100	99,70	100	**100**	100	99,40	100	**100**	100	99,40	100
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	66,50	64,40	**69,10**	66,00	66,60	**67,20**	64,60	66,70	66,10	66,70	64,60	**67,40**

7.5.4. DRD veri kümesi kullanılarak EFD yöntemine ait deneysel sonuçlar

DRD veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen eğitim ve test başarıları Çizelge 7.25'te gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.55 ile Şekil 3.60 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.25. EFD kullanılarak DRD veri kümesi üzerinde algoritmaların eğitim ve test başarıları

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	50,70	50,40	49,40	**51,30**	50,60	53,70	50,80	**54,10**	52,90	53,00	**53,70**	50,60
20	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	53,40	49,30	**54,70**	53,50	51,70	50,40	51,90	**53,20**	52,30	**53,50**	53,10	51,10
30	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	52,50	53,90	**55,20**	51,60	52,20	53,90	**56,50**	53,50	**54,90**	54,10	53,80	51,90
40	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	53,30	52,90	**58,40**	52,60	52,00	**56,10**	55,90	53,20	52,30	**52,70**	51,30	50,60
50	**Eğitim Başarısı**	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100
	Test Başarısı	57,50	56,20	**60,20**	56,60	54,90	55,70	56,30	**58,00**	52,10	56,20	540	**56,40**

DRD veri kümesi kullanılarak EFD yöntemiyle algoritmalardan elde edilen istatiksel sonuçlar Çizelge 7.26'da gösterilmiştir. Algoritmaların başarılarını karşılaştırmak amacıyla veri kümesine ait sonuçlar popülasyon sayısına bağlı olarak EK-3'de Şekil 3.61 ile Şekil 3.72 arasında grafiksel olarak gösterilmiştir.

Çizelge 7.26. EFD kullanılarak DRD veri kümesi üzerinde algoritmalardan elde edilen istatiksel sonuçlar

	İ.S.	100				200				300			
P.S.		**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**	**ABC**	**GSA**	**CSO**	**BA**
10	**Duyarlılık**	**80,00**	80,00	60,00	80,00	80,00	**100**	80,00	100	**100**	100	100	80,00
	Özgüllük	20,00	20,00	**40,00**	20,00	**20,00**	0,00	20,00	0,00	0,00	0,00	0,00	**20,00**
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
20	**Duyarlılık**	**100**	80,00	100	100	**100**	60,00	100	100	**100**	100	100	100
	Özgüllük	0,00	**20,00**	0,00	0,00	0,,00	**40,00**	0,00	0,00	**0,00**	0,00	0,00	0,00
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
30	**Duyarlılık**	**100**	90,80	95,40	80,00	**100**	94,30	93,80	100	**100**	95,00	93,9	80,00
	Özgüllük	0,00	12,40	10,20	**20,00**	0,00	7,22	**12,60**	0,00	0,00	7,82	9,42	**20,00**
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
40	**Duyarlılık**	**100**	90,80	91,40	100	60,00	85,30	85,80	**100**	**100**	92,10	93,20	100
	Özgüllük	0,00	12,50	**17,00**	0,00	**40,00**	24,30	20,70	0,00	0,00	**10,50**	6,63	0,00
	Artı Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
	Eksi Yorum Gücü	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00	**0,00**	0,00	0,00	0,00
50	**Duyarlılık**	90,80	**92,70**	86,90	87,80	**94,00**	90,60	92,80	87,60	93,70	92,70	**94,20**	92,30
	Özgüllük	18,40	13,00	**27,10**	18,90	11,50	16,30	13,20	**21,40**	7,81	**13,00**	11,50	11,60
	Artı Yorum Gücü	57,00	55,80	**59,90**	57,20	54,90	55,30	56,20	**58,30**	52,00	55,80	53,00	**56,70**
	Eksi Yorum Gücü	61,80	61,30	**62,60**	51,90	**81,10**	57,30	57,70	56,20	44,20	61,30	**71,90**	60,60

7.6. Uygulama Sonuçlarının Değerlendirilmesi

Özellik seçimi yöntemlerinde algoritmalara ait eğitim, test başarıları ve istatiksel sonuçları Çizelge 7.27'de verilmiştir.

WPBC veri kümesi için ÖS yöntemi kullanılarak elde edilen eğitim ve test sonuçları incelendiğinde deneysel çalışmalarda kullanılan tüm iterasyonlarda ve tüm popülasyon değerlerinde eğitim başarısı olarak tüm algoritmalardan %100 başarı elde edildiği görülmüştür. Test başarıları kıyaslandığında deneysel çalışmalarda kullanılan ABC, GSA, CSO ve BA algoritmalarının ikili versiyonlarının farklı iterasyon ve popülasyon değerlerinde birbirlerine üstün oldukları gözlemlenmiştir. WPBC orijinal

veri kümesindeki eğitim başarısı % 75 iken ÖS yöntemiyle elde edilen sonuçlarda dört algoritmadan elde edilen en iyi değerin orijinal veri setindekinden daha başarılı olduğu görülmüştür. İstatiksel sonuçlara bakıldığında tüm parametre değerlerinde özgüllük değerinin genellikle %100 olduğu görülmüştür. En iyi algoritmaya ait eksi yorum gücü değerinin orijinal veri setindekinden daha iyi olduğu görülmüştür. Genel olarak sonuçlar incelendiğinde, WPBC veri kümesi üzerinde CSO'nun daha iyi olduğu, BCW'de BA'nın daha başarılı olduğu görülmektedir. QSAR veri kümesi için GSA daha başarılı iken DRD'de ise BA daha başarılı sonuçlar elde etmiştir.

Çizelge 7.27. ÖS işleminde algoritmalara ait eğitim, test başarıları ve istatiksel sonuçlar

Veri Kümeleri	Algoritma	Eğitim Başarısı	Test Başarısı (Ortalama)	Test Başarısı (Standart Sapma)	Test Başarısı (Minimum)	Test Başarısı (Maksimum)	Duyarlılık	Özgüllük	Artı Yorum Gücü	Eksi Yorum Gücü
WPBC	**Salt DVM**	100	*75,00*	*3,21*	*68,00*	*77,30*	0,*00*	100	0,*00*	75,*00*
	ABC	100	*77,00*	*4,81*	*70,00*	*82,50*	38,30	85,10	28,90	87,60
	GSA	100	*80,50*	*3,26*	*77,5*	*82,50*	0,*00*	100	0,*00*	80,50
	CSO	100	***83**,00*	*3,26*	*80,00*	*87,50*	0,*00*	99,40	0,*00*	83,40
	BA	100	*80,00*	*2,50*	*77,50*	*82,50*	1,25	98,50	0,*00*	70,*00*
BCW	**Salt DVM**	100	*89,00*	*5,20*	*80,40*	*92,40*	80,47	92,41	87,72	93,96
	ABC	100	*95,90*	*0,32*	*95,70*	*96,40*	95,80	95,90	92,30	97,80
	GSA	100	*96,60*	*2,22*	*92,60*	*98,60*	95,20	97,20	94,50	97,70
	CSO	100	*96,40*	*1,01*	*95,00*	*97,14*	99,20	94,80	91,70	99,50
	BA	100	***96,70***	*1,20*	*95,71*	*98,57*	98,50	95,90	92,50	99,10
QSAR	**Salt DVM**	99,95	*73,64*	*3,15*	*70,12*	*75,23*	29,83	95,91	77,67	72,87
	ABC	99,20	*77,50*	*2,60*	*73,45*	*80,56*	53,50	89,70	72,40	79,30
	GSA	100	***80,80***	*1,30*	*79,10*	*83,00*	70,60	85,60	69,90	85,90
	CSO	100	*78,00*	*1,28*	*76,30*	*79,60*	44,20	94,40	79,50	77,80
	BA	100	*78,50*	*3,15*	*74,90*	*81,50*	51,40	91,70	74,80	79,40
DRD	**Salt DVM**	99,78	*61,81*	*1,20*	*60,30*	*63,10*	66,15	57,03	63,58	59,82
	ABC	99,60	*62,50*	*3,30*	*57,10*	*65,80*	64,40	60,20	64,80	59,90
	GSA	99,50	*67,50*	*3,74*	*61,90*	*72,20*	61,70	74,70	72,90	63,80
	CSO	99,80	*68,60*	*3,71*	*62,30*	*71,40*	69,60	50,70	61,40	54,40
	BA	99,40	***69,90***	*3,48*	*68,80*	*72,20*	70,90	66,50	70,50	67,30

EWD yöntemlerinde algoritmalara ait eğitim, test başarıları ve istatiksel sonuçlar Çizelge 7.28'de verilmiştir. Sonuçlar incelendiğinde, WPBC veri kümesi üzerinde ABC algoritmasının daha iyi performans sergilediği, BCW'de BA'nın daha başarılı olduğu görülmektedir. QSAR veri kümesi için ABC daha başarılı iken DRD'de ise GSA daha başarılı sonuçlar elde etmiştir.

Çizelge 7.28. EWD işleminde algoritmalara ait eğitim, test başarıları ve istatiksel sonuçlar

Veri Kümeleri	Algoritma	Eğitim Başarısı	Test Başarısı (Ortalama)	Test Başarısı (Standart Sapma)	Test Başarısı (Minimum)	Test Başarısı (Maksimum)	Duyarlılık	Özgüllük	Artı Yorum Gücü	Eksi Yorum Gücü
WPBC	**Salt DVM**	100	*75,00*	*4,20*	*71,20*	*78,20*	0,*00*	100	0,*00*	75,*00*
	ABC	100	***82,50***	*7,70*	*72,50*	*92,50*	15,7	99,40	0,*00*	82,60
	GSA	100	*79,00*	*2,85*	*75,00*	*82,50*	0,*00*	100	0,*00*	79,*00*
	CSO	100	*82,00*	*4,47*	*77,50*	*87,50*	0,*00*	100	0,*00*	82,*00*
	BA	100	*81,50*	*3,80*	*77,50*	*85,00*	0,*00*	100	0,*00*	81,50
BCW	**Salt DVM**	100	*89,00*	*3,40*	*83,50*	*91,42*	80,47	92,41	87,72	93,96
	ABC	100	*94,40*	*1,99*	*91,40*	*95,70*	99,6	91,80	86,10	99,80
	GSA	100	*95,00*	*1,95*	*92,10*	*97,14*	98,1	93,40	89,10	98,90
	CSO	100	*95,40*	*1,72*	*93,57*	*97,14*	86,6	100	93,10	0,*00*
	BA	100	***96,00***	*1,70*	*93,50*	*97,80*	100	93,80	89,80	100
QSAR	**Salt DVM**	99,95	*73,64*	*5,12*	*68,12*	*77,30*	29,83	95,91	77,67	72,87
	ABC	100	***77,40***	*1,70*	*75,30*	*80,00*	50,30	89,60	68,50	80,10
	GSA	98,20	*70,00*	*3,12*	*64,90*	*72,50*	15,*00*	98,10	81,*00*	69,40
	CSO	99,80	*70,60*	*4,02*	*65,80*	*76,50*	11,60	97,90	72,70	70,60
	BA	98,30	*71,40*	*1,44*	*69,70*	*73,00*	17,50	98,*00*	81,60	70,60
DRD	**Salt DVM**	99,78	*61,81*	*4,10*	*56,40*	*65,00*	66,15	57,03	63,58	59,82
	ABC	100	***65,00***	*2,00*	*62,30*	*68,30*	64,50	66,*00*	69,40	60,80
	GSA	92,60	*59,80*	*1,37*	*55,40*	*64,10*	82,20	31,*00*	57,80	59,90
	CSO	97,70	*58,60*	*1,24*	*53,20*	*64,06*	89,70	22,*00*	56,10	66,20
	BA	93,20	*59,60*	*3,18*	*58,00*	*63,20*	82,30	29,40	55,70	60,70

EFD yöntemlerinde algoritmalara ait eğitim, test başarıları ve istatiksel sonuçlar Çizelge 7.29'da verilmiştir. Sonuçlar incelendiğinde, WPBC veri kümesi üzerinde CSO

algoritmasının daha iyi performans sergilediği gözlemlenirken diğer üç veri seti üzerinde algoritmalarının başarısız olduğu görülmüştür.

Çizelge 7.29. EFD işleminde algoritmalara ait eğitim, test başarıları ve istatiksel sonuçlar

Veri Küneleri	Algoritma	Eğitim Başarısı	Test Başarısı (Ortalama)	Test Başarısı (Standart Sapma)	Test Başarısı (Minimum)	Test Başarısı (Maksimum)	Duyarlılık	Özgüllük	Arti Yorum Gücü	Eksi Yorum Gücü
WPBC	**Salt DVM**	100	*75,00*	*1,40*	*74,20*	*77,43*	0,*00*	100	0,*00*	75,*00*
	ABC	100	*81,00*	*8,02*	*70,00*	*90,00*	0,*00*	100	0,*00*	81,*00*
	GSA	100	*81,50*	*4,54*	*75,00*	*87,5*	0,*00*	100	0,*00*	81,50
	CSO	100	***83**,00*	*4,80*	*77,50*	*90,00*	0,*00*	100	0,*00*	83,*00*
	BA	100	*81,50*	*3,80*	*77,50*	*85,00*	0,*00*	100	0,*00*	81,50
BCW	**Salt DVM**	100	***89**,00*	*3,40*	*86,00*	*91,20*	80,47	92,41	87,72	93,96
	ABC	100	*68,40*	*0,44*	*65,70*	*70,00*	0,*00*	100	0,*00*	68,40
	GSA	100	*66,90*	*3,21*	*65,00*	*70,00*	0,*00*	100	0,*00*	66,90
	CSO	100	*69,00*	*7,90*	*55,70*	*75,00*	0,*00*	100	0,*00*	69,*00*
	BA	100	*67,70*	*3,20*	*63,57*	*70,00*	0,*00*	100	0,*00*	67,70
QSAR	**Salt DVM**	99,95	***73,64***	*2,10*	*70,12*	*75,00*	29,83	95,91	77,67	72,87
	ABC	100	*69,90*	*3,00*	*66,30*	*73,40*	50,30	89,60	68,60	80,10
	GSA	98,20	*70,00*	*5,30*	*64,92*	*77,25*	15,*00*	98,10	81,*00*	69,40
	CSO	99,80	*69,70*	*3,56*	*64,45*	*74,40*	11,60	97,90	72,70	70,60
	BA	98,50	*69,00*	*2,21*	*68,72*	*69,66*	12,30	97,40	71,10	69,*00*
DRD	**Salt DVM**	99,78	***61,81***	*5,23*	*55,2*	*63,00*	66,15	57,03	63,58	59,82
	ABC	100	*57,50*	*4,23*	*51,51*	*62,83*	64,50	66,*00*	69,40	60,80
	GSA	100	*56,20*	*1,27*	*54,11*	*57,57*	83,90	37,10	59,70	67,60
	CSO	100	*60,20*	*2,42*	*57,70*	*61,47*	89,70	22,*00*	56,10	66,20
	BA	100	*58,00*	*3,46*	*53,67*	*59,74*	82,30	29,40	55,70	60,70

ÖS yöntemi kullanarak BCW veri kümesinden elde edilen eğitim başarısının WPBC gibi tüm parametre değerlerinde % 100 başarıya ulaştığı görülmüştür. Test başarılarının ise orijinal BCW veri kümesinden elde edilen % 89 değerinin üzerinde olduğu tüm parametre değerlerinde gözlemlenmiş ve grafiklerde algoritmalarının başarıları gösterilmiştir. Deneysel çalışmalarda kullanılan algoritmaların tamamında,

tüm iterasyon ve popülasyon değerlerinde duyarlılık değerinin orijinal veri kümesindeki sonuçtan daha yüksek başarı kaydettiği gözlemlenmiştir. En iyi algoritma değerinden elde edilen özgüllük sonucunun salt veri kümesindeki sonuçtan daha iyi olduğu görülmüştür. Artı ve eksi yorum gücü oranlarında da durum aynı şekildedir. BCW veri setinde algoritmalarında farklı iterasyon ve popülasyon değerlerinde birbirlerine üstün oldukları gözlemlenmiştir.

QSAR veri setinde ÖS yöntemiyle elde edilen sonuçlar incelendiğinde eğitim başarısının bazı parametre değerlerinde %100'e ulaştığı görülmüştür. Diğer parametre değerlerinde ise algoritmaların %100 değerine çok yakın başarı kaydettiği gözlemlenmiştir. Algoritmaların tümünün tüm parametre değerlerinde salt veri kümesinden elde edilen % 73'lük test başarısından daha yüksek başarı elde ettiği görülmüştür. Ayrıca GSA algoritmasının ikili versiyonunun eğitim ve test oranlarında diğer algoritmalardan daha başarılı olduğu gözlemlenmiş ve GSA'nın üstünlüğü grafiksel olarak sunulmuştur. Duyarlılık oranı incelendiğinde tüm algoritmalardan elde edilen sonuçların orijinal veri kümesindeki duyarlılık sonucunun iki katından daha iyi sonuç ürettiği tüm parametre aralıklarında gözlemlenmiştir. Özgüllük oranında ise tüm parametre değerlerinde algoritmalardan elde edilen değerin başarısız olduğu gözlemlenmiştir. Bununla birlikte artı yorum gücü oranında salt veri kümesindeki değere yakın sonuçlar elde edilirken eksi yorum gücü için ise algoritmalardan daha iyi sonuçlar elde edilmiştir.

DRD veri kümesi incelendiğinde ÖS yöntemi kullanılarak algoritmalardan elde edilen eğitim sonuçlarının hiçbirinde %100 başarı elde edememesine rağmen algoritmaların salt veri kümesindeki eğitim başarı oranını bazı parametre değerlerinde geçtiği gözlemlenmiştir. Ayrıca CSO'nun eğitim aşamasında diğer algoritmalardan daha iyi performans verdiği görülmüştür. Test başarılarına bakıldığında ise parametre değerlerinin büyük çoğunluğunda algoritmaların hepsinin başarılı sonuçlar ürettiği görülmüştür. İstatiksel sonuçlara bakıldığında duyarlılık oranında bazı algoritmaların iyi sonuç vermesine rağmen genellikle duyarlılık oranının % 66 değerinin altında olduğu gözlemlenmiştir. Özgüllük oranında ise algoritmaların genel olarak başarı kaydettiği görülmüştür. Artı ve eksi yorum gücü oranlarında parametre değerlerinin büyük çoğunluğunda tüm algoritmaların daha başarılı sonuçlar elde ettiği gözlemlenmiştir.

EWD yöntemi kullanılarak WPBC veri kümesinden elde edilen eğitim sonuçlarına bakıldığında tüm iterasyon ve popülasyon değerlerinde deneysel

çalışmalarda kullanılan tüm algoritmaların eğitim başarısının %100 olduğu görülmüştür. Test başarısı incelendiğinde EWD yöntemiyle elde edilen test sonucunun bazı algoritmalarda başarılı sonuç verdiği gözlemlenmiştir. Ayrıca test başarısına göre ABC algoritmasının diğer algoritmalardan genellikle başarılı performans sergilediği gözlemlenmiş ve bu başarı grafiksel olarak gösterilmiştir. Özgüllük oranları analiz edildiğinde algoritmaların genellikle %100'e yakın başarı ürettiği görülmektedir. Eksi yorum gücü oranlarında ise algoritmaların genellikle daha iyi performans verdiği gözlemlenmiştir.

EWD yöntemiyle BCW veri kümesi kullanılarak algoritmalardan elde edilen eğitim başarılarına bakıldığında parametre değerlerinin hepsinde algoritmaların %100 başarı kaydettiği gözlemlenmiştir. Test başarısında ise parametre değerlerinin büyük çoğunluğunda algoritmaların başarılı olduğu gözlemlenmiştir. ABC algoritmasının hem test hem eğitim başarısında diğer algoritmalardan iterasyon ve popülasyon değerlerinin büyük çoğunluğunda daha başarılı olduğu görülmüş ve grafik olarak ABC'nin bu başarısı gösterilmiştir. Duyarlılık değerlerinde tüm algoritmaların genellikle başarılı olduğu görülmüştür. Bazı parametre değerlerinde duyarlılık oranının %100 olduğu gözlemlenmiştir. Özgüllük oranlarında ise algoritmaların salt veri kümesindeki özgüllük oranına yaklaşmasına rağmen algoritmaların genellikle başarısız olduğu görülmüştür. Artı yorum gücü oranında başarı kaydedilemezken eksi yorum gücü oranında ise algoritmaların farklı iterasyon ve popülasyon değerlerinde başarılı olduğu gözlemlenmiştir.

QSAR veri kümesinden EWD yöntemiyle elde edilen sonuçlar incelendiğinde genellikle eğitim başarısının iyi sonuç verdiği gözlemlenmiş olup ABC algoritmasının tüm parametre aralıklarında eğitim başarı sonucunun % 100 olduğu gözlemlenmiştir. Test başarısında ise yine ABC'nin üstünlüğü gözlemlenmiş ve algoritmanın bu başarısı grafiksel olarak sunulmuştur. Duyarlılık, özgüllük, artı ve eksi yorum gücü sonuçlarının genel başarısının farklı parametre değerlerinde iyi olduğu gözlemlenmiştir.

DBR veri kümesinden EWD yöntemi kullanılarak elde edilen eğitim başarı sonuçları incelendiğinde yine ABC'nin %100 net bir başarı sergilerken GSA'nın ise eğitim başarısının düşük olduğu gözlemlenmiştir. Test başarısında da ABC'nin tüm parametre değerlerinde diğer algoritmalara üstünlüğü görülmüştür. ABC'nin eğitim ve test başarıları grafik olarak da gösterilmiştir. Duyarlılık sonuçları incelendiğinde genellikle CSO algoritmasının daha yüksek başarı ortaya koyduğu görülmüştür. Özgüllük sonuçlarında ise ABC algoritmasının daha başarılı performans sergilediği

gözlemlenmiştir. Artı ve eksi yorum gücü sonuçları analiz edildiğinde ise genel olarak algoritmalarının farklı iterasyon ve popülasyon değerlerinde başarılı oldukları gözlemlenmiştir.

EFD yöntemi kullanılarak WPBC veri kümesi üzerinde ABC, GSA, CSO ve BA algoritmalarından elde edilen eğitim başarı sonuçları analiz edildiğinde algoritmaların hepsinin tüm parametre aralıklarında % 100 başarı ortaya koyduğu görülmüştür. Eğitim başarısında CSO algoritmasının iterasyon sayısı 100 iken diğer algoritmalardan ve salt veri kümesindeki test başarı değerinden daha yüksek performans sağladığı görülmüştür. GSA'nın ise iterasyon sayısı 300 iken diğer algoritmalardan daha yüksek bir başarı ortaya koyduğu görülmüştür. Özgüllük oranlarının deneysel çalışmalarda kullanılan tüm parametre aralıklarında % 100 olarak başarı sağladığı gözlemlenmiştir. Eksi yorum gücü oranlarında CSO ve GSA'nın diğer algoritmalara kısmen daha üstün olduğu gözlemlenmiştir.

BCW veri kümesi kullanılarak EFD yöntemiyle elde edilen eğitim sonuçlarına bakıldığında deneysel çalışmalarda kullanılan algoritmaların hepsinin tüm parametre değerlerinde % 100 şeklinde başarı kaydettiği görülmüştür. Test başarılarında ise tüm algoritmaların başarısız olduğu gözlemlenmiştir. Tüm algoritmaların parametre değerlerinin hepsinde özgüllük oranlarının % 100 başarı verdiği gözlemlenirken eksi yorum gücü oranlarında ise tüm algoritmaların başarısız olduğu gözlemlenmiştir.

EFD yöntemi kullanılarak QSAR veri kümesi üzerinde algoritmalardan elde edilen eğitim sonuçlarına bakıldığında tüm algoritmaların eğitim başarısının % 100 olduğu gözlemlenmiştir. Test başarılarında ise algoritmaların tüm parametre aralıklarında başarısız olduğu görülmüştür. Özgüllük oranlarında algoritmaların hepsinin farklı parametre aralıklarında % 100 başarı ortaya koydukları gözlemlenirken algoritmaların hepsinin duyarlılık ve eksi yorum gücü oranlarında ise başarısız olduğu görülmektedir.

DRD veri kümesinden EFD yöntemi kullanılarak elde edilen sonuçlar incelendiğinde eğitim başarı oranının tüm algoritmalar için her bir iterasyon ve popülasyon değerlerinde %100 başarı sergilediği gözlemlenmiştir. Test başarı sonuçlarında ise EFD yöntemiyle elde edilen sonuçlarda tüm algoritmaların salt veri kümesindeki test başarısının altında kaldığı görülmüştür. Duyarlılık başarısı değerlendirildiğinde tüm algoritmaların %100'e ulaşan değerler ürettiği ve her bir parametre aralığında başarılı olduğu gözlemlenmiştir. Popülasyon sayısı 50 iken elde

edilen istatiksel sonuçların diğer popülasyon sayılarındaki sonuçlardan daha başarılı olduğu gözlemlenmiştir.

8. SONUÇLAR VE ÖNERİLER

8.1. Sonuçlar

Birçok alanda karşılaşılan problemlerden biri olan sınıflandırma problemlerin çözmek ve sınıflandırma başarısını artırmak amacıyla farklı alanlarda yeni yöntemler üzerinde çalışmalar yapılmaktadır. Bu tekniklerden biri olan ve veri madenciliğinde kullanılan ÖS bu yüksek lisans tezinde kullanılmıştır. Alt küme seçimi olarak bilinen ÖS makine öğrenmesinde geniş bir kullanım alanına sahip bir yöntemdir. UCI veri kümesi kütüphanesinden alınan dört farklı veri kümesi kullanılarak ÖS'nin performansı değerlendirilmiştir. Optimizasyon algoritmalarının ikili versiyonları kullanılarak ÖS işleminin gerçekleşmesi sağlanmıştır. Optimizasyon algoritmaları yardımıyla eğitim başarısı en yüksek olacak şekilde en uygun model bulunarak elde edilen model üzerinden test işlemi uygulanmıştır. ÖS yönteminde genellikle algoritmaların eğitim başarılarının yüksek olduğu gözlemlenmiştir. Test başarılarının ise iterasyona ve popülasyon değerine göre değiştiği gözlemlenmiş olup genel olarak ise test başarısının iyi olduğu görülmüştür. ÖS işlemiyle veri kümelerine ait özellik sayısı azaltılarak hesaplama maliyeti azaltılırken buna karşın daha yüksek eğitim başarısı elde edilmiştir. Bu yöntem sayesinde işlem süresi kısaltılmış aynı zamanda veri kümesi için gerekli depolama alanı azaltmıştır.

Ayrıklaştırma, veri madenciliği ve makine öğrenmesi yöntemlerinde yaygın olarak kullanılan bir veri önişleme yöntemidir. Literatürde çok fazla ayrıklaştırma yöntemi yer almaktadır. Bu yüksek lisans tezinde denetimsiz ayrıklaştırma yöntemlerinden olan EWD ve EFD kullanılmıştır. Dört farklı optimizasyon algoritmasının sürekli versiyonu kullanılarak eğitim başarısı en yüksek olacak şekilde en uygun model elde edilip işlem sonucunda test işlemine tabi tutulmuştur. EWD ve EFD yöntemlerinden elde edilen test ve eğitim başarıları kıyas edildiğinde EWD yönteminin daha başarılı olduğu görülmüştür. Algoritmalarının ise farklı parametre değerlerinde birbirlerine üstünlükleri olduğu gözlemlenmiştir.

8.2. Öneriler

ÖS yöntemi ikili vektör uzayını girdi olarak kabul eden bir yapıya sahip olduğundan algoritmaların ikili versiyonları kullanılmıştır. ÖS yönteminin başarısını artırmak amacıyla yeni geliştirilen optimizasyon algoritmalarının sürekli versiyonları

veya algoritmalarının hibrit versiyonları ikili olarak tasarlanarak ÖS yöntemine uyarlanabilir, UCI kütüphanesindeki diğer veri setleri ya da gerçek dünya problemlerinden elde edilen veri setleri üzerinde başarıları analiz edilebilir. Bunun dışında literatürde big data olarak bilinen büyük veriler doğru analiz metotları ile yorumlandığında şirketlerin stratejik kararlarını doğru bir biçimde almalarını sağlamaktadır. Fakat bu tür verilerde veri hacminin büyük olması depolama yükünü ve hesaplama maliyetini arttırmakta ve işlem süresini uzatmaktadır. ÖS yöntemine tabi tutularak bu tür veri kümelerinden zaman açısından daha hızlı ve performans açısından daha başarılı sonuçlar elde edilebilir.

Özellikle gerçek dünya problemlerinde karşılaşılan veriler ayrık bir yapıya sahip olup bu tarz problemlerin çözümünde ayrıklaştırma yöntemine ihtiyaç duyulmaktadır. Sürekli değerleri ayrık hale getirmek için literatürde çok fazla ayrıklaştırma yöntemi yer almaktadır. Optimizasyon algoritmaları kullanılarak bu yöntemlerin başarısını artırmak amacıyla mevcut yöntemler yerine meta-sezgisel yöntemlerle ayrıklaştırma işlemi yapılabilir. Bu tez çalışmasında optimizasyon algoritmalarının sürekli versiyonu kullanılmıştır. Bunun yerine algoritmalar ayrık hale getirilerek sınıflandırma başarısında daha yüksek performans elde edilebilir.

KAYNAKLAR

Adam A, Shapiai MI, Tumari MZM, Mohamad MS, Mubin M, 2014. Feature Selection and Classifier Parameters Estimation for EEG Signals Peak Detection Using Particle Swarm Optimization. Sci World J.

Ahila R, Sadasivam V, Manimala K, 2012. Particle Swarm Optimization-Based Feature Selection and Parameter Optimization for Power System Disturbances Classification. Applied Artificial Intelligence, 26, 9, 832-61.

Akar M, Hekim M, Orhan U, 2015. Mechanical fault detection in permanent magnet synchronous motors using equal width discretization-based probability distribution and a neural network model. Turk J Electr Eng Co, 23, 3, 813-23.

Akay B, 2009. Nümerik Optimizasyon Problemlerinde Yapay Arı Kolonisi (Artificial Bee Colony) Algoritmasının Performans Analizi. Erciyes Üniversitesi Fen Bilimleri Enstitüsü, 301s.

Akyol S, Alataş B, 2012. Güncel sürü zekası algoritmaları. Nevşehir Üniversitesi Fen Bilimleri Enstitü Dergisi 1, 36-50.

Aneesh MU, Masand AAK, Manikantan K, 2012. Optimal Feature Selection based on Image Pre-processing using Accelerated Binary Particle Swarm Optimization for Enhanced Face Recognition. International Conference on Communication Technology and System Design 2011, 30, 750-8.

Arnulf B, Graf A, Borer S, 2001. Normalization in Support Vector Machines. Lecture Notes in Computer Science, 2191:, 277-82.

Aydin I, Karakose M, Akin E, 2011. A multi-objective artificial immune algorithm for parameter optimization in support vector machine. Applied Soft Computing, 11, 1, 120-9.

Banka H, Dara S, 2015. A Hamming distance based binary particle swarm optimization (HDBPSO) algorithm for high dimensional feature selection, classification and validation. Pattern Recogn Lett, 52, 94-100.

Bao Y, Hu Z, Xiong T, 2013. A PSO and pattern search based memetic algorithm for SVMs parameters optimization. Neurocomputing, 117, 98-106.

Bereta M, Burczyński T, 2007. Comparing binary and real-valued coding in hybrid immune algorithm for feature selection and classification of ECG signals. Engineering Applications of Artificial Intelligence, 20, 5, 571-85.

Bhadra T, Bandyopadhyay S, Maulik U, 2012. Differential Evolution Based Optimization of SVM Parameters for Meta Classifier Design. Procedia Technology, 4, 50-7.

Biba M, Esposito F, Ferilli S, Mauro ND, Basile TMA, 2007. Unsupervised Discretization Using Kernel Density Estimation. Proceedings of the 20th international joint conference on Artifical intelligence (IJCAI), 6-12 Ocak 2007 Hyderabad India, SanFrancisco, CA, USA:Morgan Kaufmann Publishers Inc., 696-701.

Boubezoul A, Paris S, 2012. Application of global optimization methods to model and feature selection. Pattern Recognition, 45, 10, 3676-86.

Boulle M, 2005. Optimal Bin Number for Equal Frequency Discretization in Supervized Learning. Intell Data Anal, 9, 2, 175-88.

Bruni R, Bianchi G, 2015. Effective Classification Using a Small Training Set Based on Discretization and Statistical Analysis. Ieee T Knowl Data En, 27, 9, 2349-61.

Cachopo AC, Oliveira AL, 2006. Empirical Evaluation of Centroid–based Models for Single–label Text Categorization. IESC–ID Technical Report, 1–9

Carrizosa E, Martín-Barragán B, Romero Morales D, 2014. A nested heuristic for parameter tuning in Support Vector Machines. Computers & Operations Research, 43, 328-34.

Cervante L, Xue B, Zhang MJ, Shang L, 2012. Binary Particle Swarm Optimisation for Feature Selection: A Filter Based Approach. Ieee C Evol Computat.

Chakrabarti S, Cox E, Frank E, Gting RH, Han J, Jiang X, Kamber M, Lightstone SD, Nadeau TP, Neapolitan RE, Pyle D, Refaat M, Schneider M, Teorey TJ, Witten IH, 2008. Data Mining Know it all. Morgan Kaufmann Publishers Inc., San Francisco, CA, 978-0-12-374629-0.

Chang Y-C, Lai P-C, Lee M-T, 2007. An integrated approach for operational knowledge acquisition of refuse incinerators. Expert Systems with Applications, 33, 2, 413-9.

Chaves R, Ramirez J, Gorriz JM, Neuroimaging AsD, 2013. Integrating discretization and association rule-based classification for Alzheimer's disease diagnosis. Expert Systems with Applications, 40, 5, 1571-8.

Che J, 2013. Support vector regression based on optimal training subset and adaptive particle swarm optimization algorithm. Applied Soft Computing, 13, 8, 3473-81.

Chen J-S, Cheng C-H, 2008. Extracting classification rule of software diagnosis using modified MEPA. Expert Systems with Applications, 34, 1, 411-8.

Chen J, Zhang C, Xue X, Liu C-L, 2013. Fast instance selection for speeding up support vector machines. Knowledge-Based Systems, 45, 1-7.

Chen L, Chen BL, Chen YX, 2011. Image Feature Selection Based on Ant Colony Optimization. Lect Notes Artif Int, 7106, 580-9.

Chen Y, Yang Y, Liu C, Li C, Li L, 2015. A hybrid application algorithm based on the support vector machine and artificial intelligence: An example of electric load forecasting. Applied Mathematical Modelling, 39, 9, 2617-32.

Cheng W-Y, Juang C-F, 2011. An incremental support vector machine-trained TS-type fuzzy system for online classification problems. Fuzzy Sets and Systems, 163, 1, 24-44.

Choi H, Yeo D, Kwon S, Kim Y, 2011. Gene selection and prediction for cancer classification using support vector machines with a reject option. Computational Statistics & Data Analysis, 55, 5, 1897-908.

Chou J-S, Cheng M-Y, Wu Y-W, Pham A-D, 2014. Optimizing parameters of support vector machine using fast messy genetic algorithm for dispute classification. Expert Systems with Applications, 41, 8, 3955-64.

Chuang LY, Chang HW, Tu CJ, Yang CH, 2008. Improved binary PSO for feature selection using gene expression data. Comput Biol Chem, 32, 1, 29-37.

Chuang LY, Tsai SW, Yang CH, 2011. Improved binary particle swarm optimization using catfish effect for feature selection. Expert Systems with Applications, 38, 10, 12699-707.

Chuang LY, Yang CH, Li JC, 2011. Chaotic maps based on binary particle swarm optimization for feature selection. Applied Soft Computing, 11, 1, 239-48.

Çamlica H, Disci R, 2008. Tanı testlerinde sınır değerlerinin belirlenmesi. Türk Radyasyon Onkolojisi 23, 1, 26-33.

Das K, Vyas OP, 2010. A Suitability Study of Discretization Methods for Associative Classifiers. International Journal of Computer Applications (0975 – 8887) 5, 10, 46-51.

De la Hoz E, De la Hoz E, Ortiz A, Ortega J, Martinez-Alvarez A, 2014. Feature selection by multi-objective optimisation: Application to network anomaly

detection by hierarchical self-organising maps. Knowledge-Based Systems, 71, 322-38.

Dehuai Z, Yuan L, Lianbo J, Li L, Gang X, 2012. Wick Sintered Temperature Forecasting Based on Support Vector Machines with Simulated Annealing. Physics Procedia, 25, 427-34.

Devos O, Downey G, Duponchel L, 2014. Simultaneous data pre-processing and SVM classification model selection based on a parallel genetic algorithm applied to spectroscopic data of olive oils. Food Chem, 148, 124-30.

Doğan Z, 2012. Ayrıklaştırma yöntemleri ve yapay sinir ağı kullanarak asenkron motorlarda arıza teşhisi. Marmara Üniversitesi, Fen Bilimleri Fakültesi.

Dougherty C, Kohavı R, Sahamı M, 1995. Supervised and Unsupervised Discretization of Continuous Features. In Armand: Prieditis, A., Russell, S.J. (ed), Machine Learning: Proceedings of the Twelfth International Conference on Machine Learning, 9-12 Temmuz 1995 Tahoe City, California, USA, Morgan Kaufmann, 194-202.

Durmaz O, 2011 Metin Sınıflandırmada Boyut Azaltmanın Etkisi ve Özellik Seçimi.

Emary E, Zawbaa HM, Grosan C, Hassenian AE, 2015. Feature Subset Selection Approach by Gray-Wolf Optimization. Adv Intell Syst, 334, 1-13.

Erguzel TT, Ozekes S, Gultekin S, Tarhan N, 2014. Ant Colony Optimization Based Feature Selection Method for QEEG Data Classification. Psychiat Invest, 11, 3, 243-50.

Feng W, Zhang Q, Hu G, Huang JX, 2014. Mining network data for intrusion detection through combining SVMs with ant colony networks. Future Generation Computer Systems, 37, 127-40.

Fister I, Fister D, Yang X-S, 2013. A hybrid bat algorithm. Elektrotehniski vestnik, 80(1-2), 1-7.

Forsati R, Moayedikia A, Jensen R, Shamsfard M, Meybodi MR, 2014. Enriched ant colony optimization and its application in feature selection. Neurocomputing, 142, 354-71.

Gang R, Zhuping Z, 2011. Traffic safety forecasting method by particle swarm optimization and support vector machine. Expert Systems with Applications, 38, 8, 10420-4.

García Nieto PJ, García-Gonzalo E, Alonso Fernández JR, Díaz Muñiz C, 2016. A hybrid PSO optimized SVM-based model for predicting a successful growth cycle of the Spirulina platensis from raceway experiments data. Journal of Computational and Applied Mathematics, 291, 293-303.

García Nieto PJ, García-Gonzalo E, Sánchez Lasheras F, de Cos Juez FJ, 2015. Hybrid PSO–SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. Reliability Engineering & System Safety, 138, 219-31.

Garcia R, Paraiso EC, Nievola JC, 2011. Multiobjective Optimization of Indexes Obtained by Clustering for Feature Selection Methods Evaluation in Genes Expression Microarrays. Lect Notes Comput Sc, 6936, 353-60.

Ghamisi P, Benediktsson JA, 2014. Feature selection of hyperspectral data by considering the integration of Genetic Algorithms and Particle Swarm Optimization. Image and Signal Processing for Remote Sensing Xx, 9244.

Ghamisi P, Benediktsson JA, 2015. Feature Selection Based on Hybridization of Genetic Algorithm and Particle Swarm Optimization. Ieee Geosci Remote S, 12, 2, 309-13.

Gruter C, Farina WM, 2009. The honeybee waggle dance: can we follow the steps? Trends Ecol Evol, 24, 5, 242-7.

Grzymala-Busse JW, Stefanowski J, 2001. Three Discretization Methods for Rule Induction. Int J Intell Syst, 16, 29-38.

Guo X, Li D, Zhang A, 2012. Improved Support Vector Machine Oil Price Forecast Model Based on Genetic Algorithm Optimization Parameters. AASRI Procedia, 1, 525-30.

Güraksın GE, Haklı H, Uğuz H, 2014. Support vector machines classification based on particle swarm optimization for bone age determination. Applied Soft Computing, 24, 597-602.

Hacıbeyoglu M, 2012. Bilgi sistemlerinde fark fonksiyonu tabanlı özellik seçme yönteminin geliştirilmesi. Selcuk Universitesi, Fen Bilimleri.

Han H, Wang HJ, Dong XC, 2011. Transformer Fault Dignosis Based on Feature Selection and Parameter Optimization. Enrgy Proced, 12.

Han J, Kamber M, 2006. Data Mining: Concepts and Techniques. Morgan Kaufmann, San Francisco, 311–5, 48–50

Han X, Chang X, Quan L, Xiong X, Li J, Zhang Z, Liu Y, 2014. Feature subset selection by gravitational search algorithm optimization. Information Sciences, 281, 128-46.

Han XH, Chang XM, Quan L, Xiong XY, Li JX, Zhang ZX, Liu Y, 2014. Feature subset selection by gravitational search algorithm optimization. Information Sciences, 281, 128-46.

Harish N, Mandal S, Rao S, Patil SG, 2015. Particle Swarm Optimization based support vector machine for damage level prediction of non-reshaped berm breakwater. Applied Soft Computing, 27, 313-21.

Hsieh C-T, Hu C-S, 2014. Fingerprint Recognition by Multi-objective Optimization PSO Hybrid with SVM. Journal of Applied Research and Technology, 12, 6, 1014-24.

Hsieh T-J, Hsiao H-F, Yeh W-C, 2012. Mining financial distress trend data using penalty guided support vector machines based on hybrid of particle swarm optimization and artificial bee colony algorithm. Neurocomputing, 82, 196-206.

Hsu C-N, Huang H-J, Wong T-T, 2003. Implications of the Dirichlet Assumption for Discretization of Continuous Variables in Naive Bayesian Classifiers. Machine Learning, Cilt 53, No 3, 235–63.

Hu Z, Bao Y, Xiong T, 2014. Comprehensive learning particle swarm optimization based memetic algorithm for model selection in short-term load forecasting using support vector regression. Applied Soft Computing, 25, 15-25.

Huang C-F, 2012. A hybrid stock selection model using genetic algorithms and support vector regression. Applied Soft Computing, 12, 2, 807-18.

Ilhan I, Tezel G, 2013. A genetic algorithm-support vector machine method with parameter optimization for selecting the tag SNPs. J Biomed Inform, 46, 2, 328-40.

Inderscience, 2010. Cuckoo designs spring. Alphagalileo.org, Retrieved 2010-05-27.

Jedrzejowicz P, Skakovski A, 2014. Island-based differential evolution algorithm for the discrete-continuous scheduling with continuous resource discretisation. Knowledge-Based and Intelligent Information & Engineering Systems 18th Annual Conference, Kes-2014, 35, 111-7.

Jiang S-y, Li X, Zheng Q, Wang L-x, 2009. Approximate Equal Frequency Discretization Method. 514-8.

Jiang SY, Li X, Zheng Q, Wang LX, 2009. Approximate Equal Frequency Discretization Method. Proceedings of the 2009 Wri Global Congress on Intelligent Systems, Vol Iii, 514-8.

Jin R, Breitbart Y, Muoh C, 2007. Data Discretization Unification. 183-92.

Karaboga D, 2005. An Idea Based On Honey Bee Swarm For Numerical Optimization. Technical Report-TR06, Erciyes University Engineering Faculty Computer Engineering Department, Kayseri.

Karaboga D, Basturk B, 2007. Artificial Bee Colony (ABC) Optimization Algorithm for Solving Constrained Optimization Problems. Advances in Soft Computing: Foundations of Fuzzy Logic and Soft Computing, 4529, 789–98,.

Kavzoğlu T, Çölkesen İ, 2010. Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. Harita Dergisi, 144, 7, 73-82.

Kennedy J, Eberhart R, 1997. A discrete binary version of the particle swarm algorithm. IEEE international conference on

computational cybernetics and simulation, pp 4104–8.

Khan A, Revett K, 2004. Data Mining the PIMA Dataset Using Rough Set Theory with a Special Emphasis on Rule Reduction. Proceedings of INMIC 8th International Multitopic Conference, 24-26 Aralık 2004 Lahore Pakistan, IEEE Conference Publications, 334-9.

Koçoğlu FÖ, 2012. Tez-Veri madenciliği sürecinde veri ayrıklaştırma yöntemlerinin karşılaştırılması ve bir uygulama. İstanbul Üniversitesi Fen Bilimleri Enstitüsü.

Koçoğlu FÖ, 2012. Veri Madenciliği Sürecinde Veri Ayriklaştirma Yöntemlerinin Karşilaştirilmasi Ve Bir Uygulama. İstanbul Üniversitesi Fen Bilimleri Enstitüsü.

Kotsiantis S, Kanellopoulos D, 2006. Discretization Techniques: A recent survey. GESTS International Transactions on Computer Science and Engineering, 32, 2, 47-58.

Kumar MA, Gopal M, 2009. Least squares twin support vector machines for pattern classification. Expert Systems with Applications, 36, 4, 7535-43.

Kuo RJ, Chen CM, Liao TW, Tien FC, 2013. Hybrid of artificial immune system and particle swarm optimization-based support vector machine for Radio Frequency Identification-based positioning system. Computers & Industrial Engineering, 64, 1, 333-41.

Kurgan L, Cios KJ, 2004. CAIM Discretization Algorithm. Ieee T Knowl Data En, 16, 2, 145-53.

Lee SM, Kim DS, Park JS, 2011. Cost-Sensitive Spam Detection Using Parameters Optimization and Feature Selection. J Univers Comput Sci, 17, 6, 944-60.

Lin KC, Chen SY, Hung JC, 2015. Feature Selection and Parameter Optimization of Support Vector Machines Based on Modified Artificial Fish Swarm Algorithms. Math Probl Eng.

Lin KC, Huang YH, Hung JC, Lin YT, 2015. Feature Selection and Parameter Optimization of Support Vector Machines Based on Modified Cat Swarm Optimization. Int J Distrib Sens N.

Liu B, Tian MH, Zhang CH, Li XT, 2015. Discrete Biogeography Based Optimization for Feature Selection in Molecular Signatures. Mol Inform, 34, 4, 197-215.

Liu H, Hussain F, Tan C, Dash M, 2002. Discretization: An Enabling Technique. Data Mining and Knowledge Discovery, 6, 4, 393-423.

Liu H, Shi XM, Guo DM, Zhao ZW, Yimin, 2015. Feature Selection Combined with Neural Network Structure Optimization for HIV-1 Protease Cleavage Site Prediction. Biomed Res Int.

Liu S, Tai H, Ding Q, Li D, Xu L, Wei Y, 2013. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. Mathematical and Computer Modelling, 58, 3-4, 458-65.

Liu Y, Wang G, Chen H, Dong H, Zhu X, Wang S, 2011. An Improved Particle Swarm Optimization for Feature Selection. Journal of Bionic Engineering, 8, 2, 191-200.

Maldonado S, Pérez J, Weber R, Labbé M, 2014. Feature selection for Support Vector Machines via Mixed Integer Linear Programming. Information Sciences, 279, 163-75.

Manimala K, Selvi K, Ahila R, 2011. Hybrid soft computing techniques for feature selection and parameter optimization in power quality data mining. Applied Soft Computing, 11, 8, 5485-97.

Mehta S, Parthasarathy S, Yang H, 2005. Toward Unsupervised Correlation Preserving Discretization. Ieee T Knowl Data En, 17, 9, 1174-85.

Miranda PBC, Prudêncio RBC, de Carvalho APLF, Soares C, 2014. A hybrid meta-learning architecture for multi-objective optimization of SVM parameters. Neurocomputing, 143, 27-43.

Mirjalili S, Mirjalili SM, Yang X-S, 2013. Binary bat algorithm. Neural Computing and Applications, 25, 3-4, 663-81.

Moayedikia A, Jensen R, Wiil UK, Forsati R, 2015. Weighted bee colony algorithm for discrete optimization problems with application to feature selection. Engineering Applications of Artificial Intelligence, 44, 153-67.

Moradi P, Rostami M, 2015. Integration of graph clustering with ant colony optimization for feature selection. Knowledge-Based Systems, 84, 144-61.

Nikitidis S, Nikolaidis N, Pitas I, 2012. Multiplicative update rules for incremental training of multiclass support vector machines. Pattern Recognition, 45, 5, 1838-52.

Olfati E, Zarabadipour H, Shoorehdeli MA, 2014. Feature Subset Selection and Parameters Optimization for Support Vector Machine in Breast Cancer Diagnosis. 2014 Iranian Conference on Intelligent Systems (Icis).

Olson DL, Delen D, 2008. Advanced Data Mining Techniques. Springer, Verlag Berlin Heidelberg, 978-3-540-76916-3.

Onnia V, Tico M, Saarinen J, 2001. Feature selection method using neural network. Proceeding 2001 International Conferencee on Image Processing,, Thessaloniki, Yunanistan.

Orhan U, Hekim M, Ozer M, 2012. Epileptic Seizure Detection Using Probability Distribution Based On Equal Frequency Discretization. J Med Syst, 36, 4, 2219-24.

Osuna EE, Freund R, Girosi F, 1997. Support Vector Machines: Training and Applications. Massachusetts Institute of Technology and Artificial Intelligence Laboratory, Massachusetts., A.I. Memo No. 1602, C.B.C.L. Paper No. 144.

Ozdemir S, 2014. Yerine geri koyarak örnekleme, ayrıklaştırma ve öznitelik seçme kullanan karar ağacı temelli saldırı tespit sistemi. B.S., Telecommunication Engineering, Istanbul Technical University, 2011.

Pal A, Maiti J, 2010. Development of a hybrid methodology for dimensionality reduction in Mahalanobis–Taguchi system using Mahalanobis distance and

binary particle swarm optimization. Expert Systems with Applications, 37, 2, 1286-93.

Pan H, Xia SY, Jin LZ, Xia LZ, 2011. Combining heterogeneous features for face detection using multiscale feature selection with binary particle swarm optimization. Proc Spie, 8006.

Peng X, Wang Y, Xu D, 2013. Structural twin parametric-margin support vector machine for binary classification. Knowledge-Based Systems, 49, 63-72.

Rajabioun R, 2011. Cuckoo Optimization Algorithm. Applied Soft Computing, 11, 8, 5508-18.

Rashedi E, Nezamabadi-pour H, Saryazdi S, 2009. GSA: A Gravitational Search Algorithm. Information Sciences, 179, 13, 2232-48.

Rashedi E, Nezamabadi-pour H, Saryazdi S, 2010. BGSA: binary gravitational search algorithm. Natural Computing, 9, 3, 727-45.

Rubio G, Pomares H, Rojas I, Herrera LJ, 2011. A heuristic method for parameter selection in LS-SVM: Application to time series prediction. International Journal of Forecasting, 27, 3, 725-39.

Saha S, Ekbal A, Alok AK, Spandana R, 2014. Feature selection and semi-supervised clustering using multiobjective optimization. Springerplus, 3.

Sahu B, Mishra D, 2012. A Novel Feature Selection Algorithm using Particle Swarm Optimization for Cancer Microarray Data. Procedia Engineering, 38, 27-31.

Sarac E, Ozel SA, 2014. An Ant Colony Optimization Based Feature Selection for Web Page Classification. Sci World J.

Sartakhti JS, Zangooei MH, Mozafari K, 2012. Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA). Comput Methods Programs Biomed, 108, 2, 570-9.

Sayın KE, 2013. KAK veri setinde sınıflama başarısını arttırmak için öznitelik seçme ve ayrıklaştırma. Kadir Has University, Graduate School Of Scıence And Engıneerıng.

Shima K, Todoriki M, Suzuki A, 2004. SVM-based feature selection of latent semantic features. Pattern Recogn Lett, 25, 9, 1051-7.

Shirin N, Hodais S, Vafaei JM, 2012. A Binary Model on the Basis of Cuckoo Search Algorithm. 2012 International Conference on System Engineering and Modeling (ICSEM 2012), IPCSIT vol. 34 (2012) © (2012) IACSIT Press, Singapore.

Shokri S, Sadeghi MT, Marvast MA, 2014. High reliability estimation of product quality using support vector regression and hybrid meta-heuristic algorithms. Journal of the Taiwan Institute of Chemical Engineers, 45, 5, 2225-32.

Song CY, Jiang JQ, Wu CG, Liang YC, 2011. Feature Selection Algorithm Based on Least Squares Support Vector Machine and Particle Swarm Optimization. Lect Notes Comput Sc, 6729, 275-82.

Srinivasa KG, Venugopal KR, Patnaik LM, 2007. A self-adaptive migration model genetic algorithm for data mining applications. Information Sciences, 177, 20, 4295-313.

Sun LG, de Visser CC, Chu QP, Mulder JA, 2014. A novel online adaptive kernel method with kernel centers determined by a support vector regression approach. Neurocomputing, 124, 111-9.

Tan PN, Steinbach M, Kumar V, 2006. Introduction to Data Mining. Pearson Education Inc., Boston, 0-321-32136-7.

Tasgetiren M, Suganthan P, Pan Q, 2007. A Discrete Particle Swarm Optimization Algorithm for the Generalized Traveling Salesman Problem. In: 9th annual

conference on genetic and evolutionary computation (GECCO ‘07), New York, NY, USA, 2007, pp 158–67.

Tereshko V, Loengarov A, 2005. Collective decision making in honey-bee foraging dynamics. Computing and Informaton Systems, 9(3), 1-7.

Vieira SM, Mendonça LF, Farinha GJ, Sousa JMC, 2013. Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. Applied Soft Computing, 13, 8, 3494-504.

Wang F-K, Du T, 2014. Implementing support vector regression with differential evolution to forecast motherboard shipments. Expert Systems with Applications, 41, 8, 3850-5.

Wang F, Yang Y, Lv XC, Xu J, Li L, 2014. Feature Selection using Feature Ranking, Correlation Analysis and Chaotic Binary Particle Swarm Optimization. 2014 5th Ieee International Conference on Software Engineering and Service Science (Icsess), 305-9.

Wang H, Jing XJ, Niu B, 2014. A Weighted Bacterial Colony Optimization for Feature Selection. Lect N Bioinformat, 8590, 379-89.

Wang S, Meng B, 2011. Parameter Selection Algorithm for Support Vector Machine. Procedia Environmental Sciences, 11, 538-44.

Wang X, Yang J, Teng X, Xia W, Jensen R, 2007. Feature selection based on rough sets and particle swarm optimization. Pattern Recogn Lett, 28, 4, 459-71.

Wei L, Hanning C, 2012. BABC: A Binary Version of Artificial Bee Colony Algorithm for Discrete Optimization. International Journal of Advancements in Computing Technology, 4, 14, 307-14.

Wong AKC, Chiu DKY, 1987. Synthesizing Statistical Knowledge from Incomplete Mixed-Mode Data. Ieee T Pattern Anal, PAMI-9, 6, 796-805.

Wu JY, Qiu TR, Wang L, Huang HQ, 2011. An Approach to Feature Selection Based on Ant Colony Optimization and Rough Set. Comm Com Inf Sc, 134, I, 466-71.

Wu T-H, Chang C-C, Chung S-H, 2008. A simulated annealing algorithm for manufacturing cell formation problems. Expert Systems with Applications, 34, 3, 1609-17.

Xi J, Ouyang WM, 2000. Clustering based Algorithm for Best Discretizing Continuous Valued Attributes Mini-micro systems, Cilt 21, No 10, 1025–7.

Xue B, Qin AK, Zhang MJ, 2014. An Archive Based Particle Swarm Optimisation for Feature Selection in Classification. 2014 Ieee Congress on Evolutionary Computation (Cec), 3119-26.

Yang X-S, 2012. Bat Algorithm for Multi-objective Optimisation. Int. J. Bio-Inspired Computation, Vol. 3, No. 5, 267-74.

Yang X-S, Gandomi AH, 2012. Bat Algorithm: A Novel Approach for Global Engineering Optimization.pdf. Engineering Computations, 29(5), 464 - 83.

Yang XS, 2010. A New Metaheuristic Bat-Inspired Algorithm (NICSO 2010). Springer Berlin Heidelberg, v.284, pp. 65-74.

Yang XS, Deb S, 2009. Cuckoo Search via Levy Flights. World Congress on Nature & Biologically Inspired Computing (NaBIC 2009), 210–4.

Yang Y, Webb GI, 2002. A Comparative Study of Discretization Methods for Naïve-Bayes Classifiers. İçinde: Yamaguchi, T., Hoffmann, A., Motoda, H., Compton, P. (ed.), Proceedings of the 2002 Pacific Rim Knowledge Acquisition Workshop (PKAW'02) Tokyo, Japan, Tokyo: Japanese Society for Artificial Intelligence, Tokyo, Japan, Tokyo: Japanese Society for Artificial Intelligence, 159-73.

Yang Y, Webb GI, 2005. Discretization for Data Mining. İçinde: Wang, J. (ed.), Encyclopedia of Data Warehousing and Mining, IGI-Global Disseminator of Knowledge, USA, 392-6.

Yu Z, Wong H-S, Wen G, 2011. A modified support vector machine and its application to image segmentation. Image and Vision Computing, 29, 1, 29-40.

Zeng X-P, Li Y-M, Qin J, 2009. A dynamic chain-like agent genetic algorithm for global numerical optimization and feature selection. Neurocomputing, 72, 4-6, 1214-28.

Zhai S, Jiang T, 2014. A novel particle swarm optimization trained support vector machine for automatic sense-through-foliage target recognition system. Knowledge-Based Systems, 65, 50-9.

Zhao MY, Fu C, Ji LP, Tang K, Zhou MT, 2011. Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes. Expert Systems with Applications, 38, 5, 5197-204.

Zheng H, Zhou Y, He S, Ouyang X, 2012. A Discrete Cuckoo Search Algorithm for Solving Knapsack Problems. INTERNATIONAL JOURNAL ON Advances in Information Sciences and Service Sciences, 4, 18, 331-9.

Zhou H, Pei Zhao J, Gang Zheng L, Lin Wang C, Fa Cen K, 2012. Modeling NOx emissions from coal-fired utility boilers using support vector regression with ant colony optimization. Engineering Applications of Artificial Intelligence, 25, 1, 147-58.

Zhou J, Li X, Shi X, 2012. Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. Safety Science, 50, 4, 629-44.

Zhou W, Dickerson JA, 2014. A novel class dependent feature selection method for cancer biomarker discovery. Comput Biol Med, 47, 66-75.

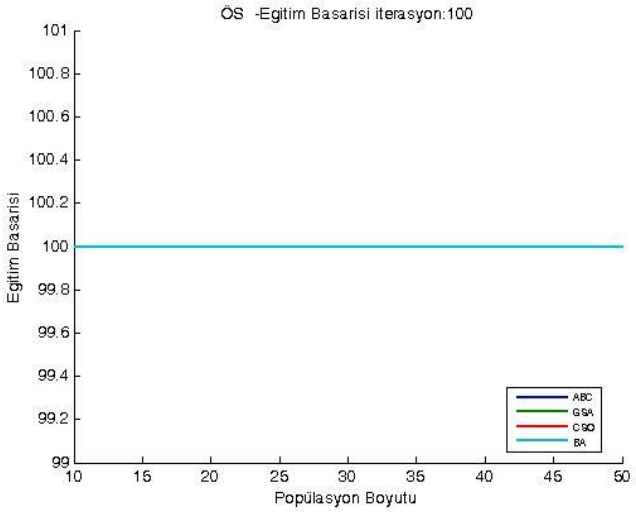
Zongker D, Jain A, 1996. Algorithms for feature selection: An evaluation. Proceedings of ICPR 96.

Online: Inderscience (27 May 2010). "Cuckoo designs spring". Alphagalileo.org. Retrieved 2010-05-27 [Ziyaret Tarihi: 4 Ekim 2015].

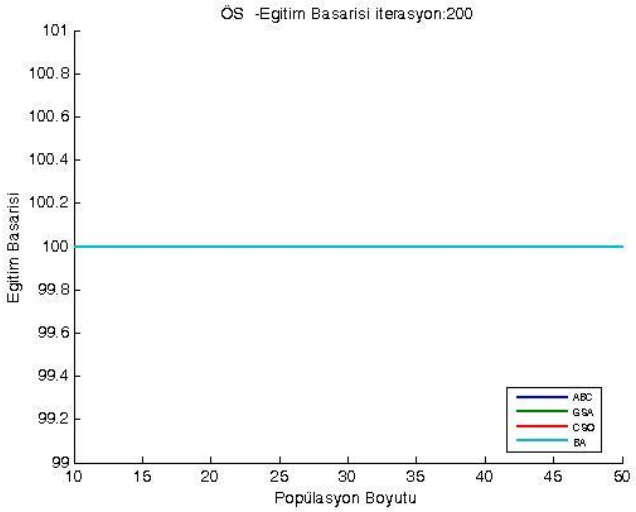
Online: Hsu, C.W., Chang, C.C., Lin, C.J., 2010, A Practical Guide to Support Vector Classification,http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf. [Ziyaret Tarihi: 4 Ekim 2015].

EKLER

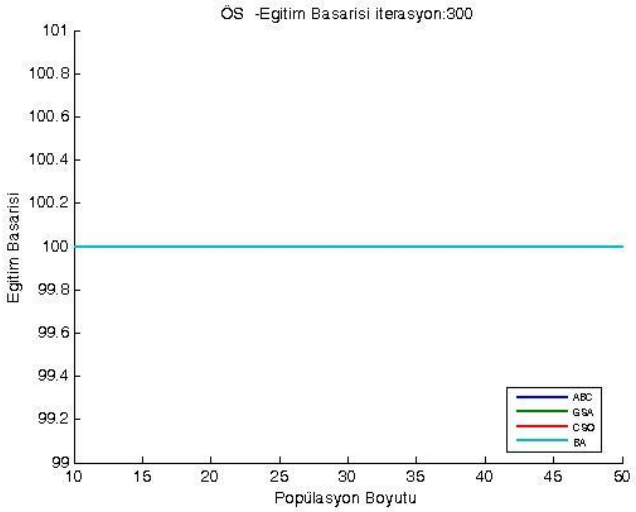
EK-1 ÖS Yöntemi İçin Test Sonuçları



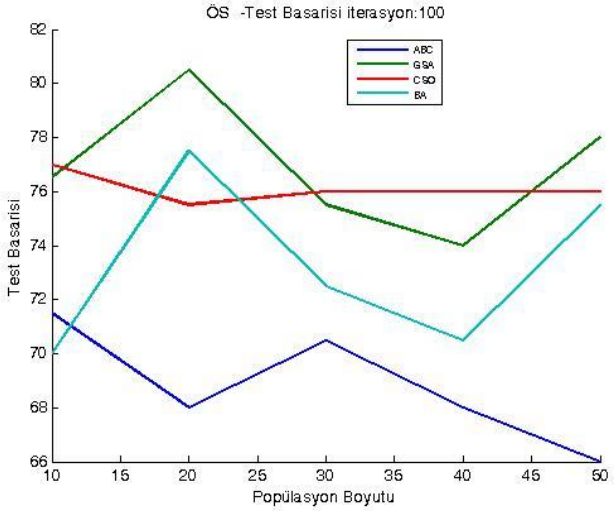
Şekil 1.1. WPBC veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



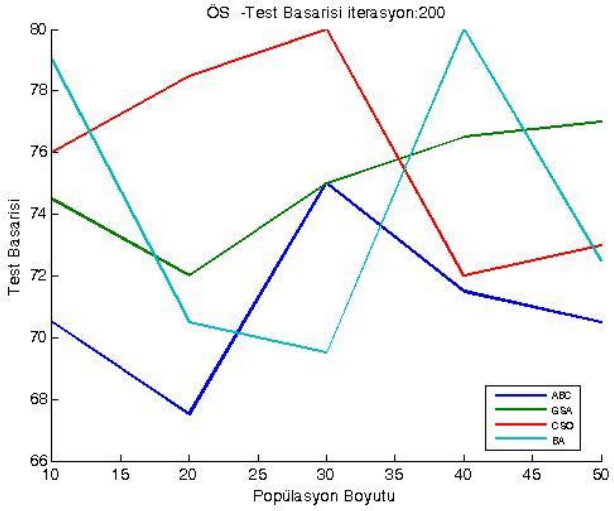
Şekil 1.2. WPBC veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



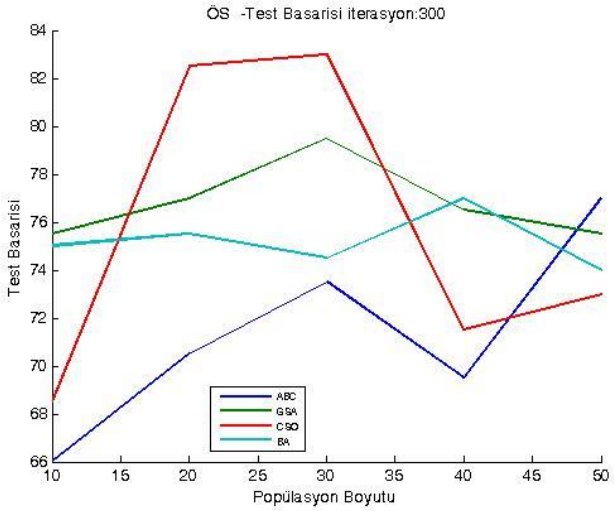
Şekil 1.3. WPBC veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



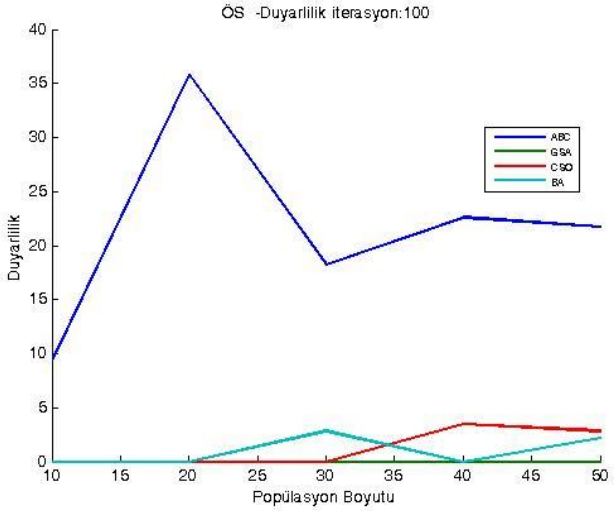
Şekil 1.4. WPBC veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



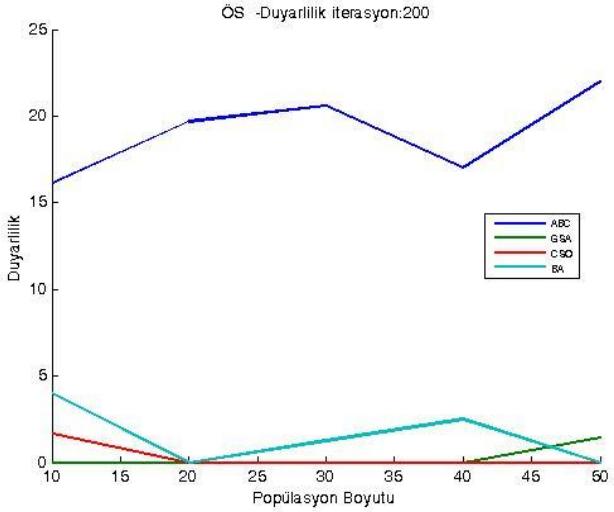
Şekil 1.5. WPBC veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



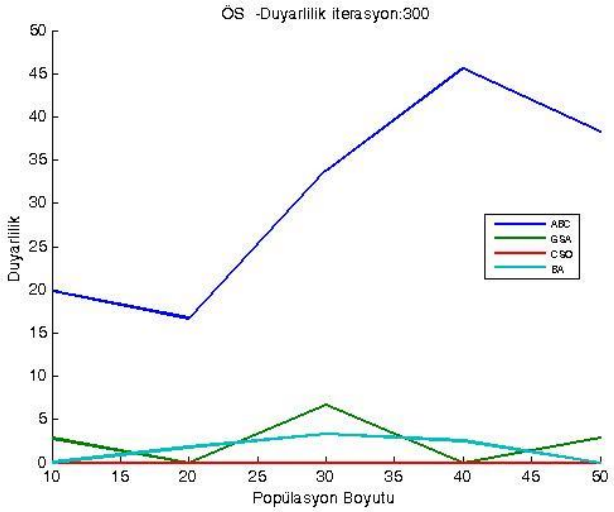
Şekil 1.6. WPBC veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



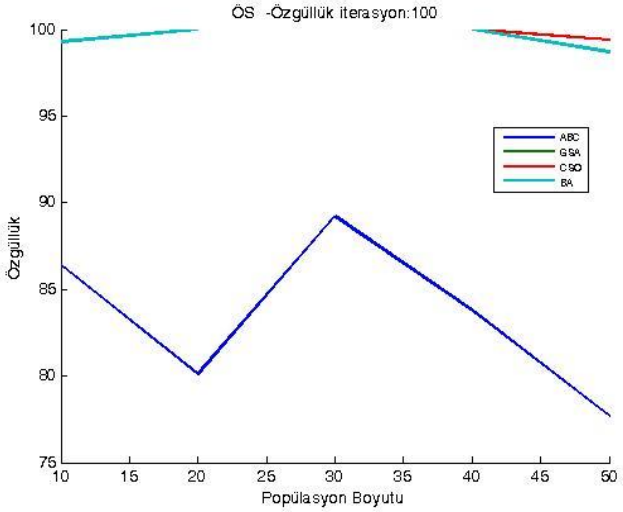
Şekil 1.7. WPBC veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



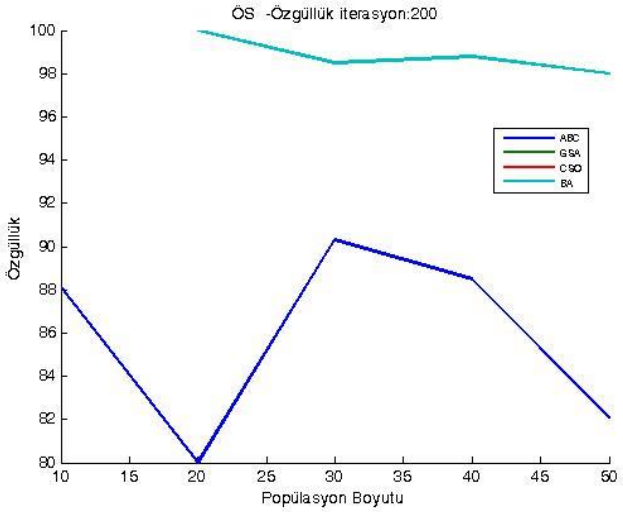
Şekil 1.8. WPBC veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



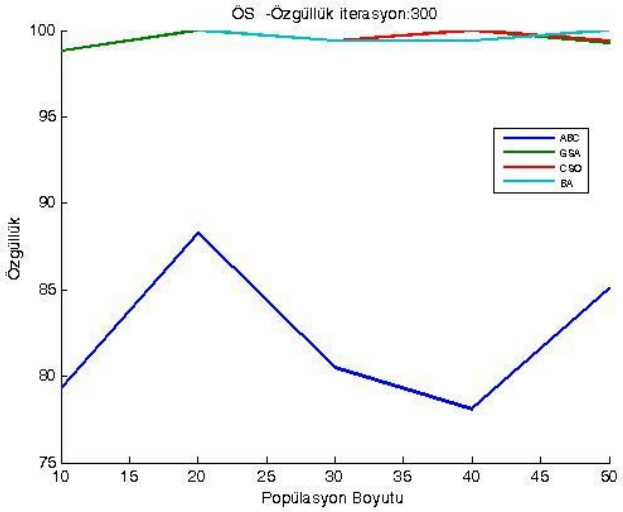
Şekil 1.9. WPBC veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



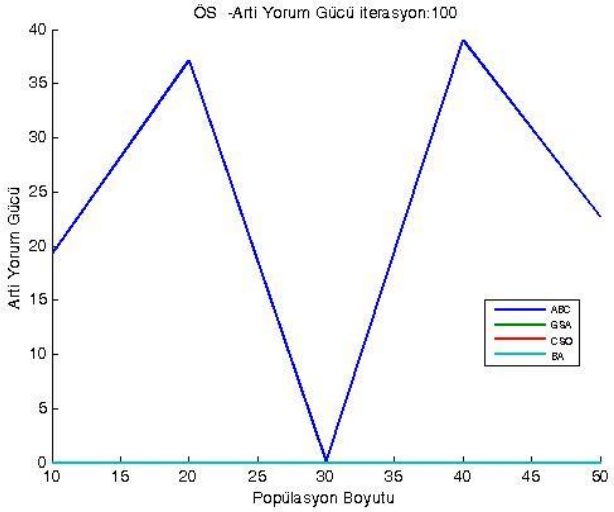
Şekil 1.10. WPBC veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



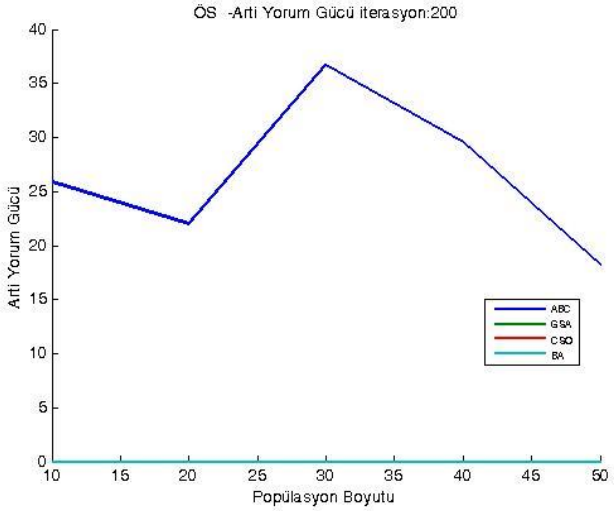
Şekil 1.11. WPBC veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



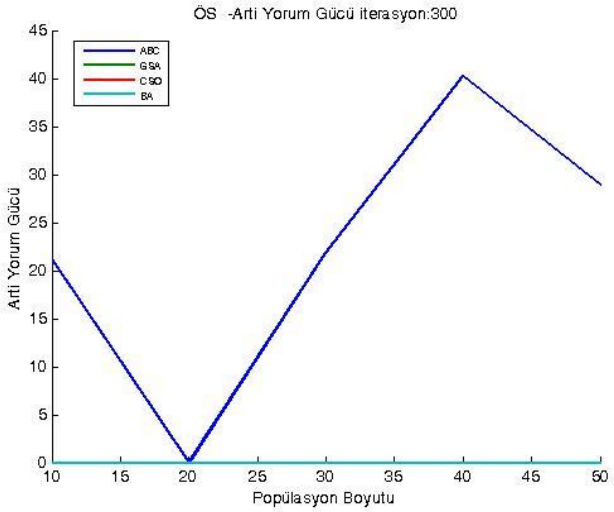
Şekil 1.12. WPBC veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



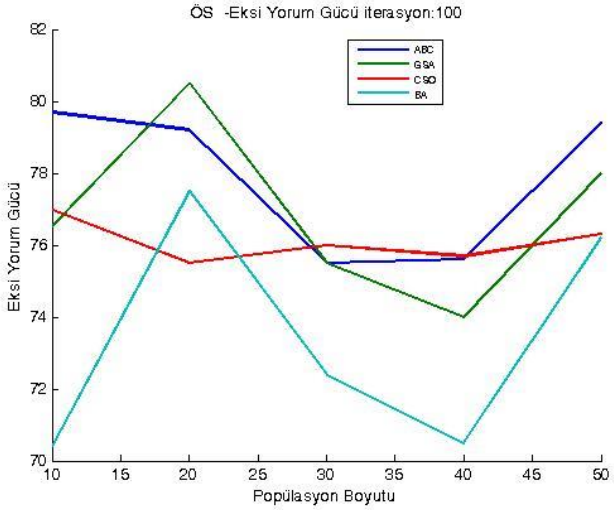
Şekil 1.13. WPBC veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



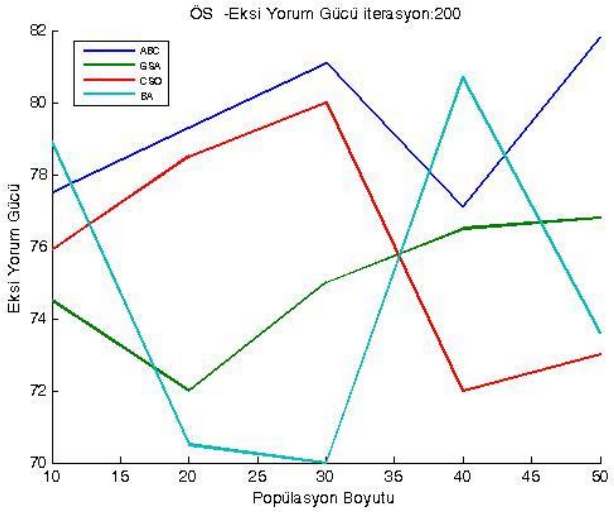
Şekil 1.14. WPBC veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



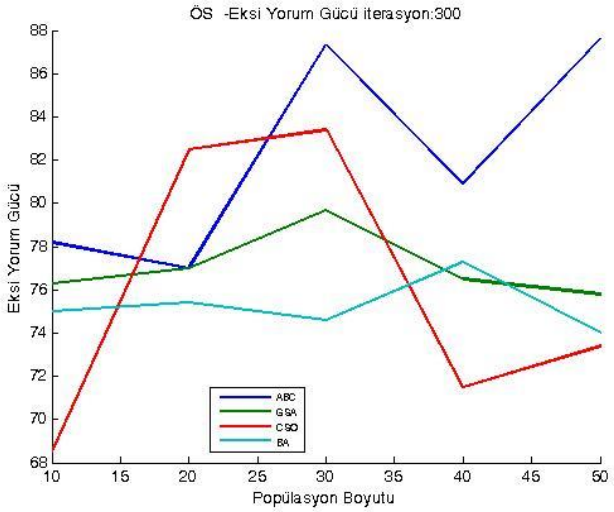
Şekil 1.15. WPBC veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



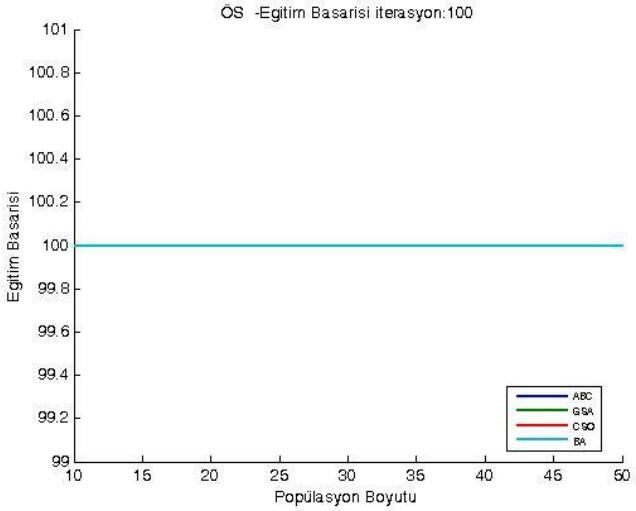
Şekil 1.16. WPBC veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



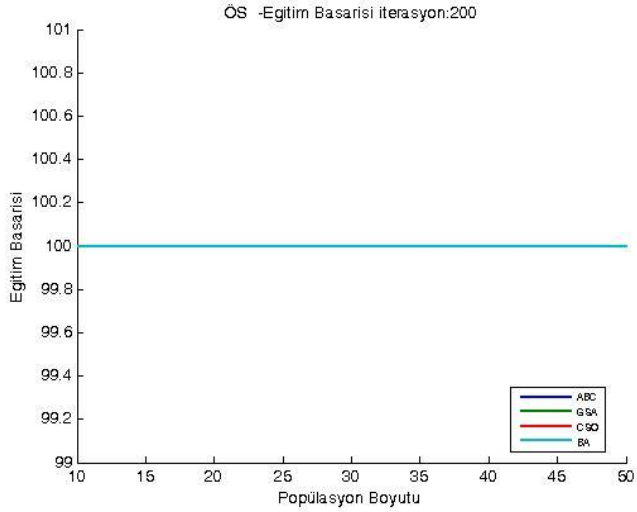
Şekil 1.17. WPBC veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



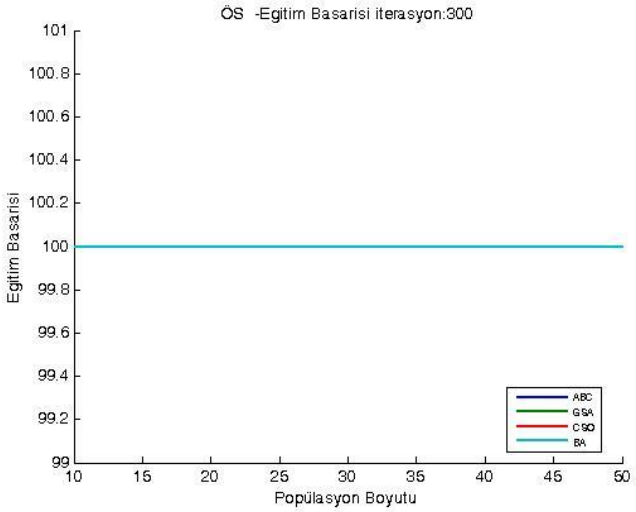
Şekil 1.18. WPBC veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



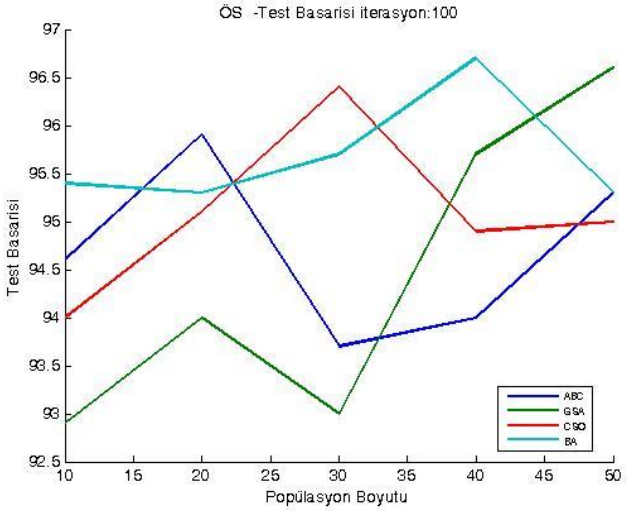
Şekil 1.19. BCW veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



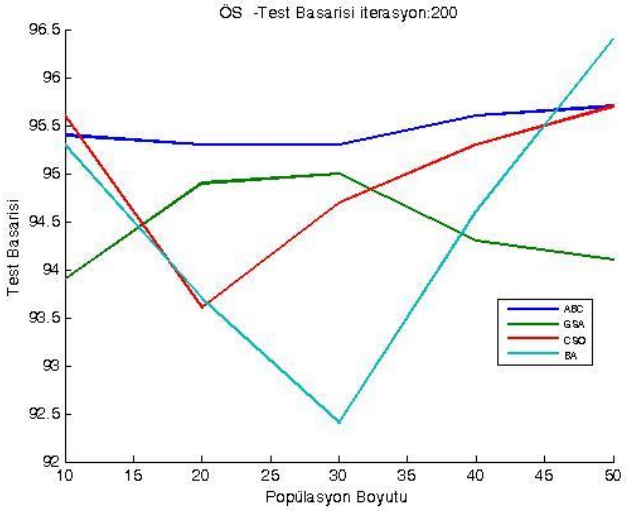
Şekil 1.20. BCW veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



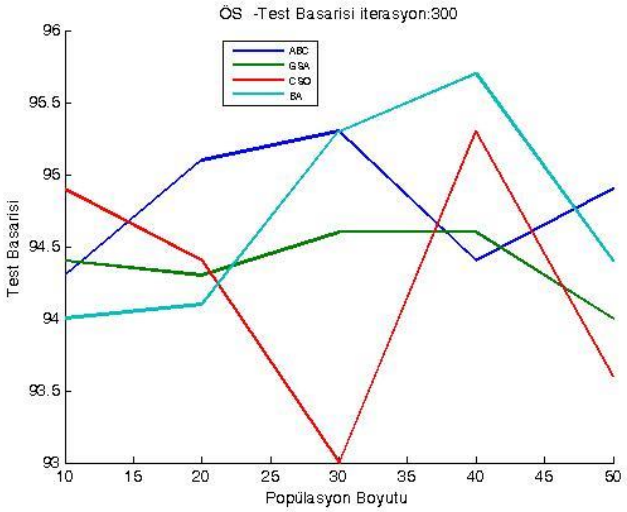
Şekil 1.21. BCW veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



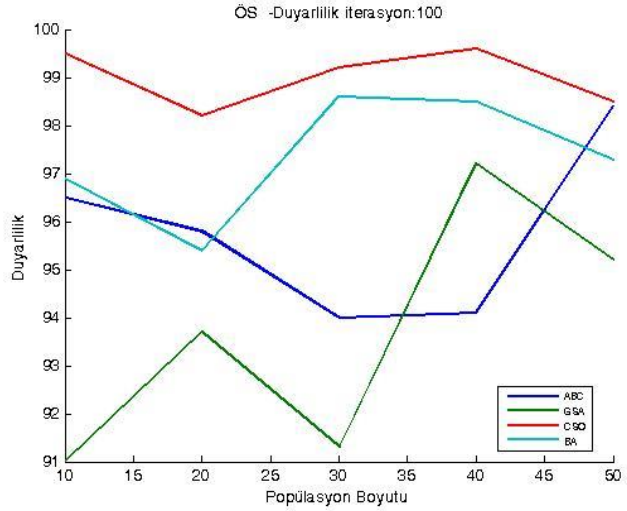
Şekil 1.22. BCW veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



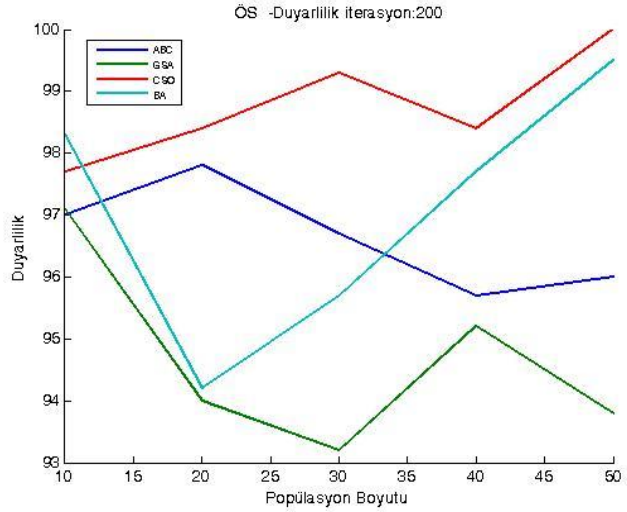
Şekil 1.23. BCW veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



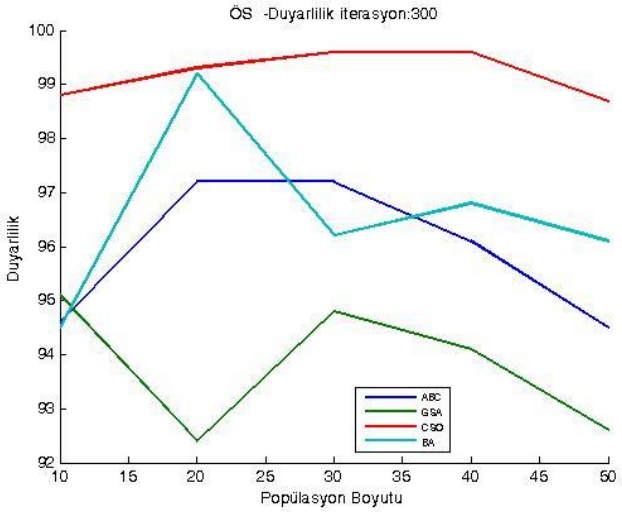
Şekil 1.24. BCW veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



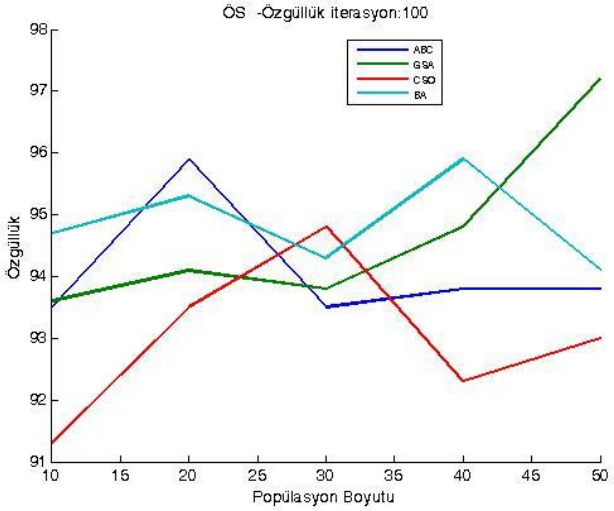
Şekil 1.25. BCW veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



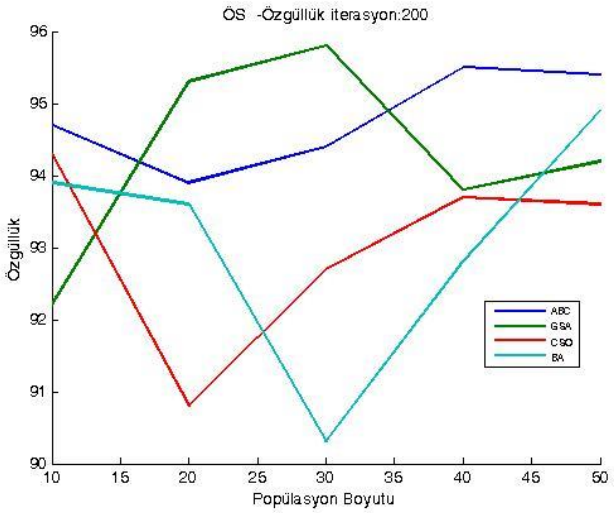
Şekil 1.26. BCW veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



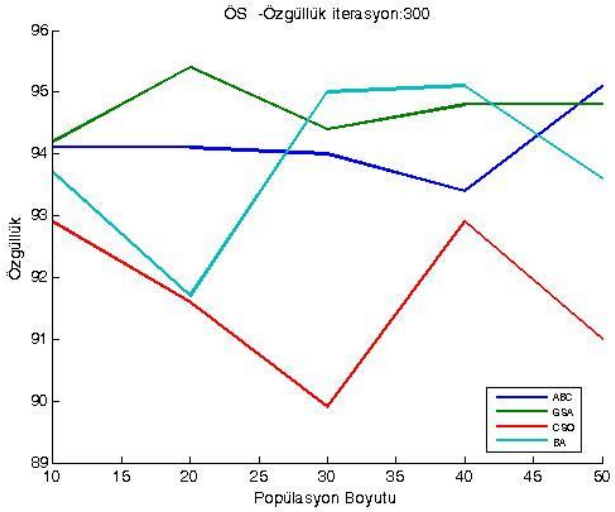
Şekil 1.27. BCW veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



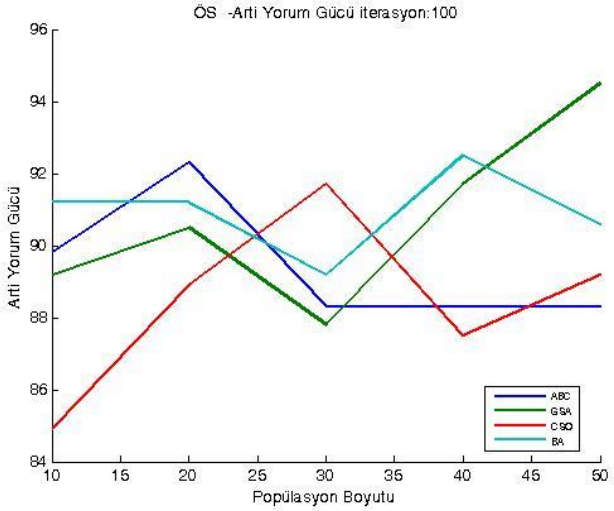
Şekil 1.28. BCW veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



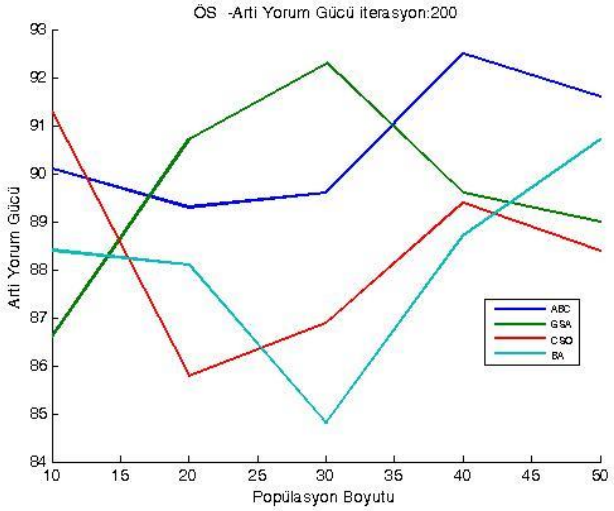
Şekil 1.29. BCW veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



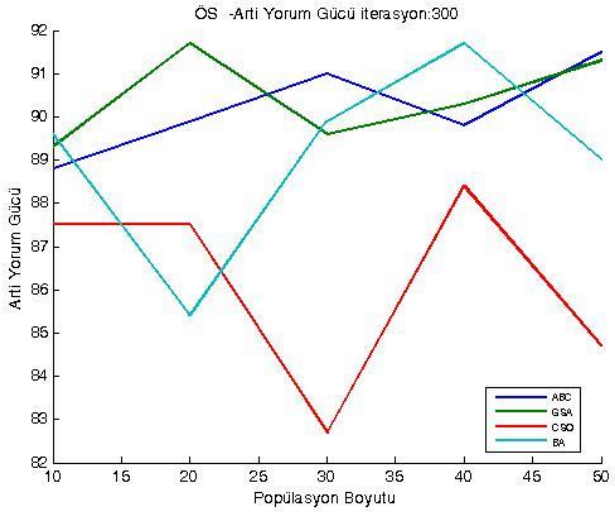
Şekil 1.30. BCW veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



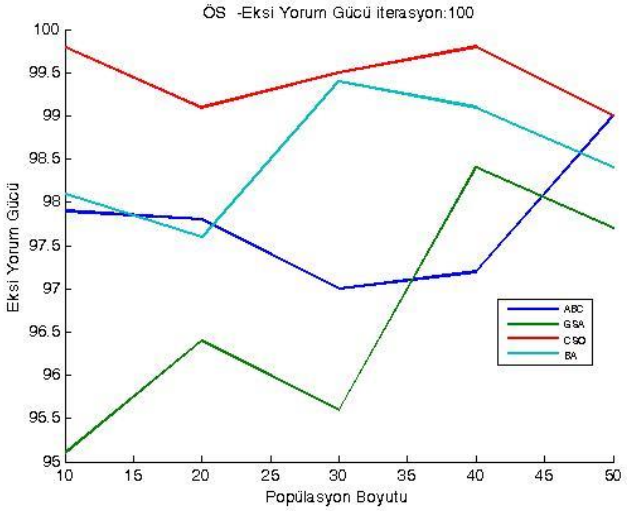
Şekil 1.31. BCW veri kümesi üzerinde 100 iterasyon için artı yorum gücü sonuçları



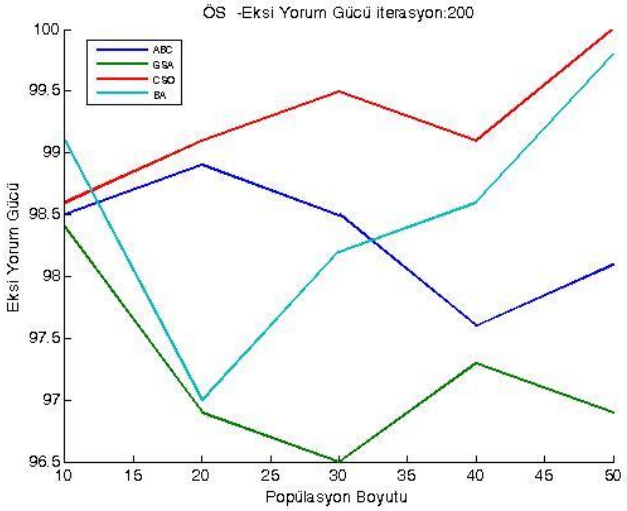
Şekil 1.32. BCW veri kümesi üzerinde 200 iterasyon için artı yorum gücü sonuçları



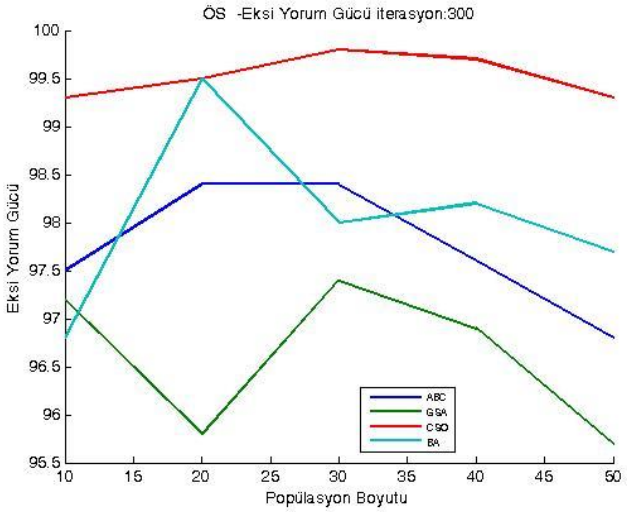
Şekil 1.33. BCW veri kümesi üzerinde 300 iterasyon için artı yorum gücü sonuçları



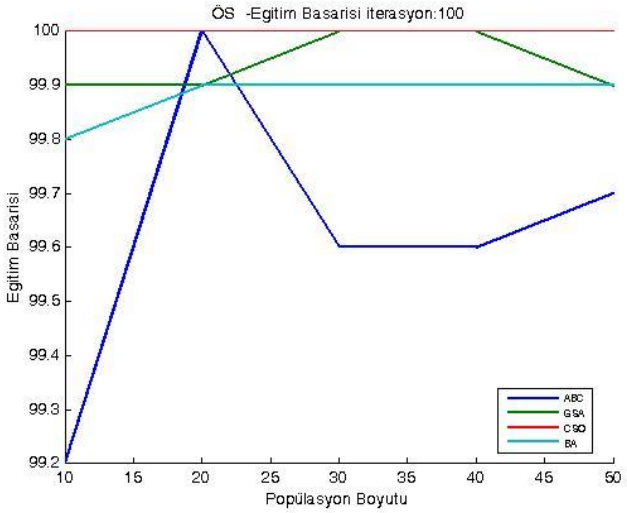
Şekil 1.34. BCW veri kümesi üzerinde 100 iterasyon için eksi yorum gücü sonuçları



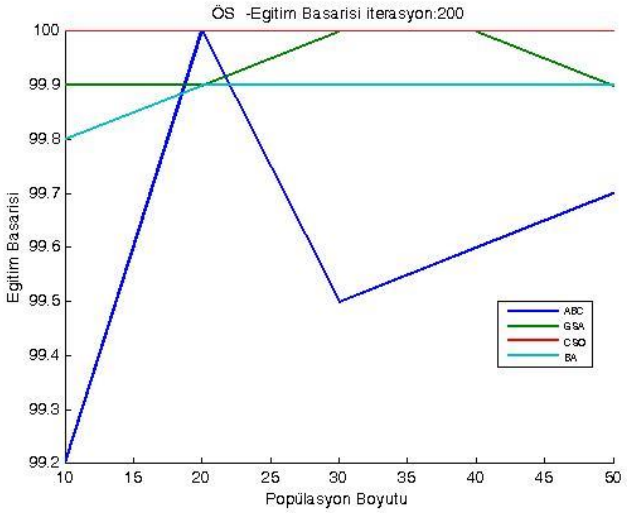
Şekil 1.35. BCW veri kümesi üzerinde 200 iterasyon için eksi yorum gücü sonuçları



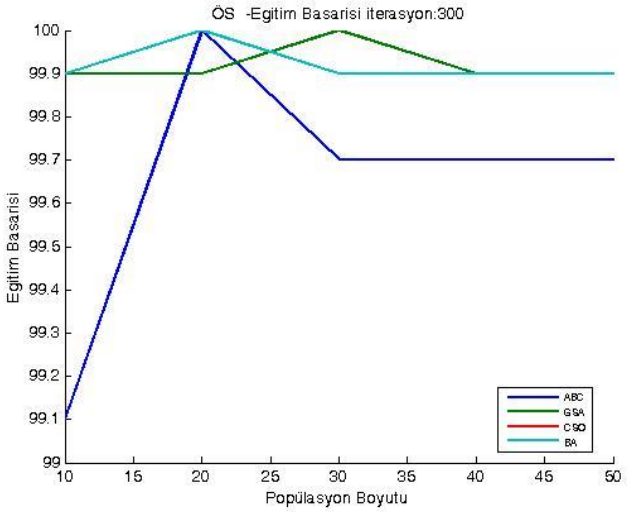
Şekil 1.36. BCW veri kümesi üzerinde 300 iterasyon için eksi yorum gücü sonuçları



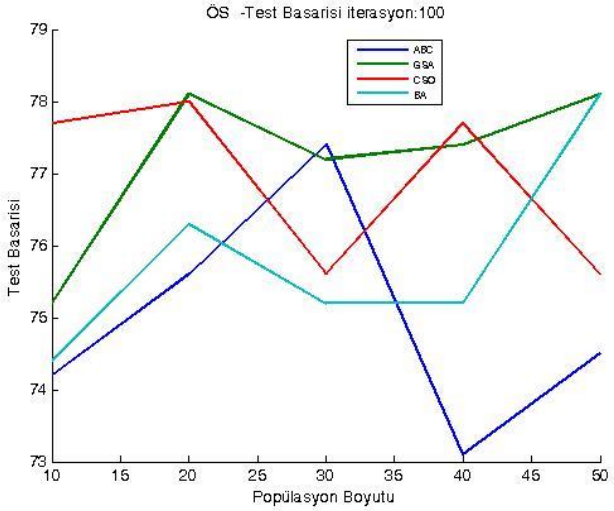
Şekil 1.37. QSAR veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



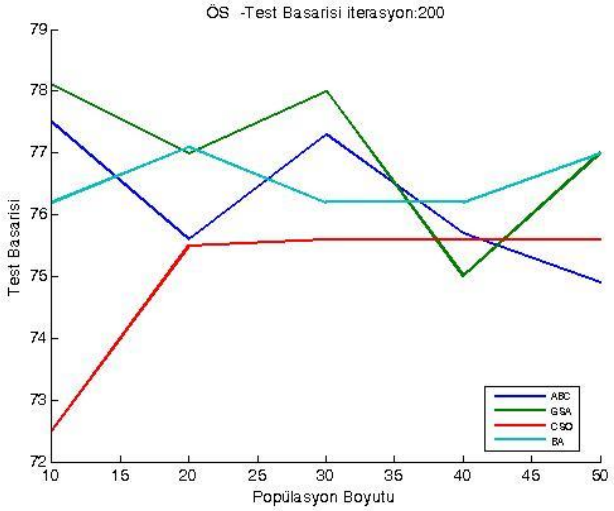
Şekil 1.38. QSAR veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



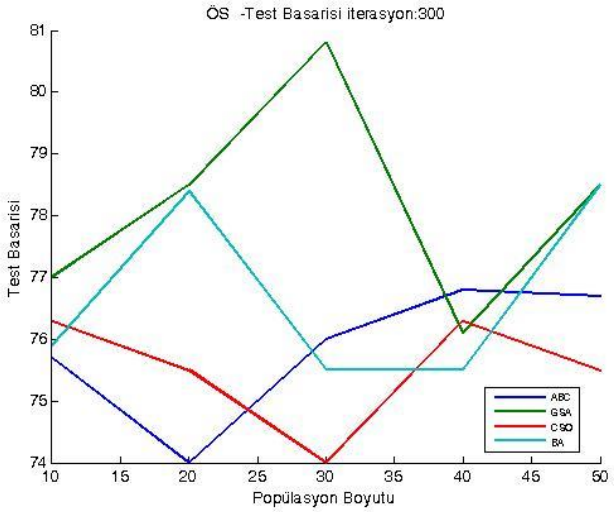
Şekil 1.39. QSAR veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



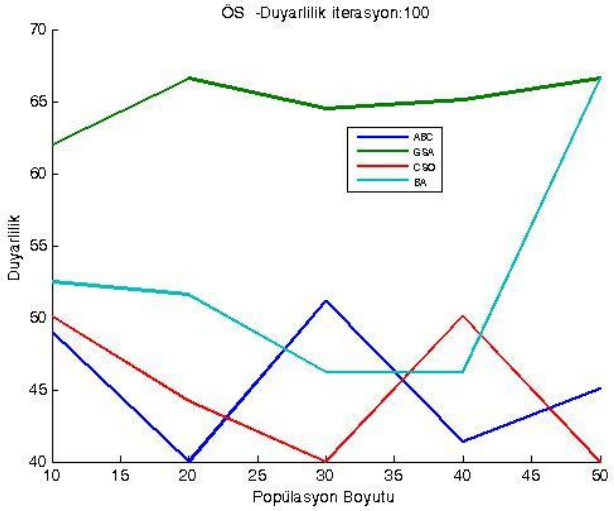
Şekil 1.40. QSAR veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



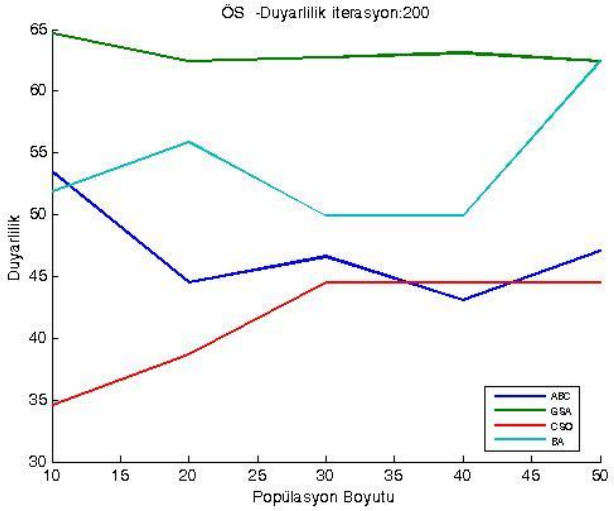
Şekil 1.41. QSAR veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



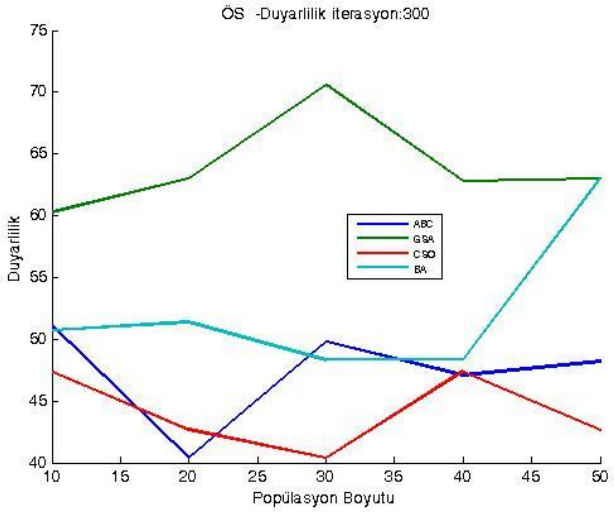
Şekil 1.42. QSAR veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



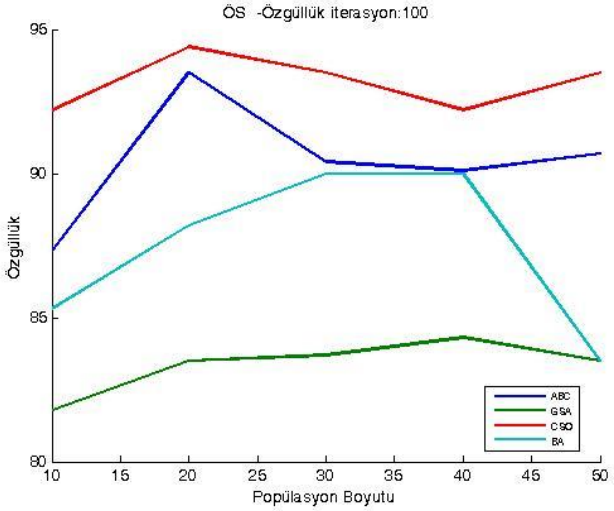
Şekil 1.43. QSAR veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



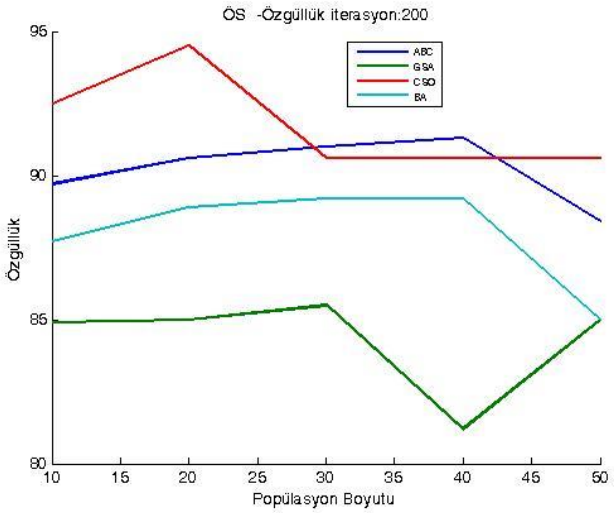
Şekil 1.44. QSAR veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



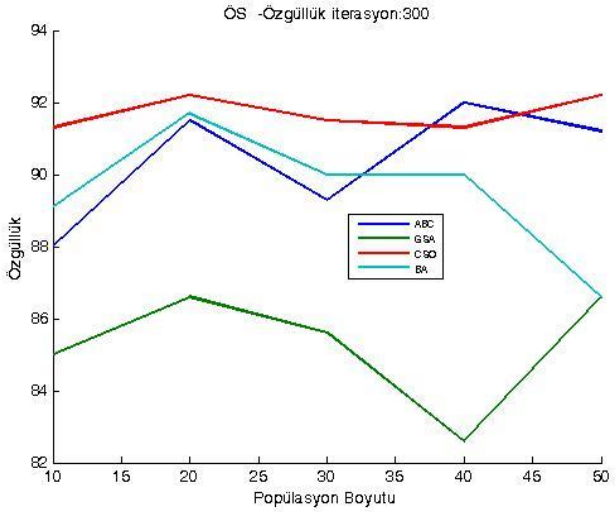
Şekil 1.45. QSAR veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



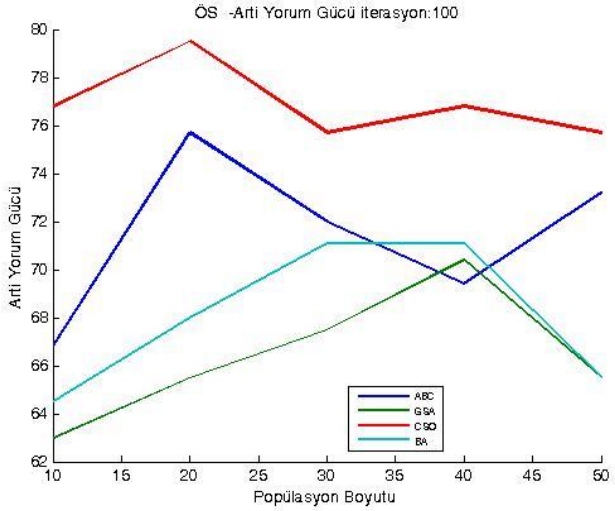
Şekil 1.46. QSAR veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



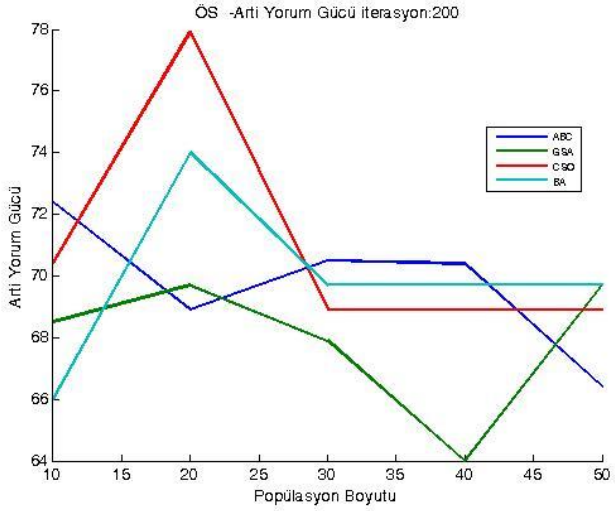
Şekil 1.47. QSAR veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



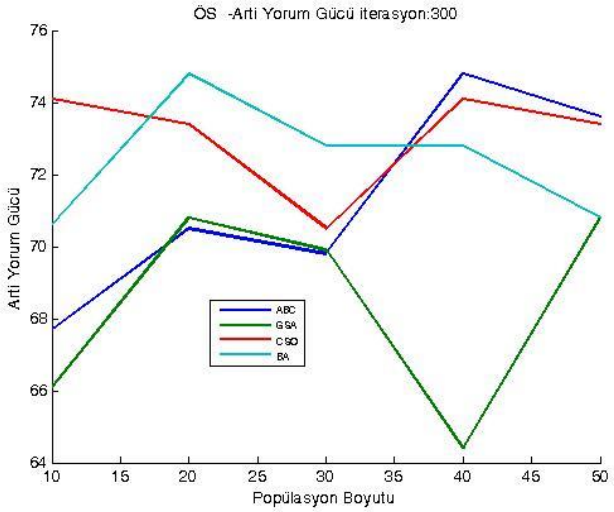
Şekil 1.48. QSAR veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



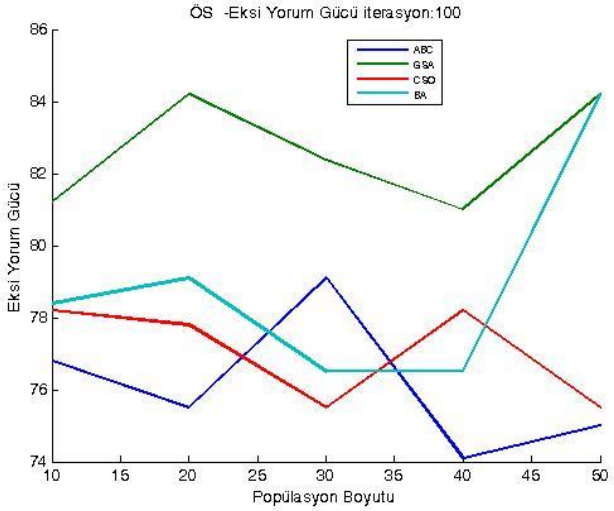
Şekil 1.49. QSAR veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



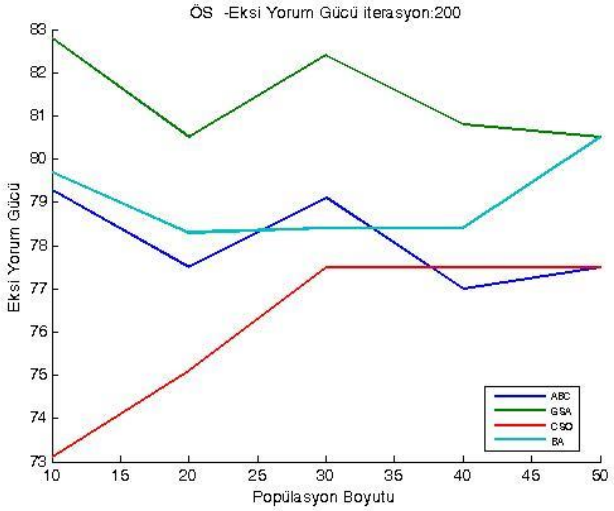
Şekil 1.50. QSAR veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



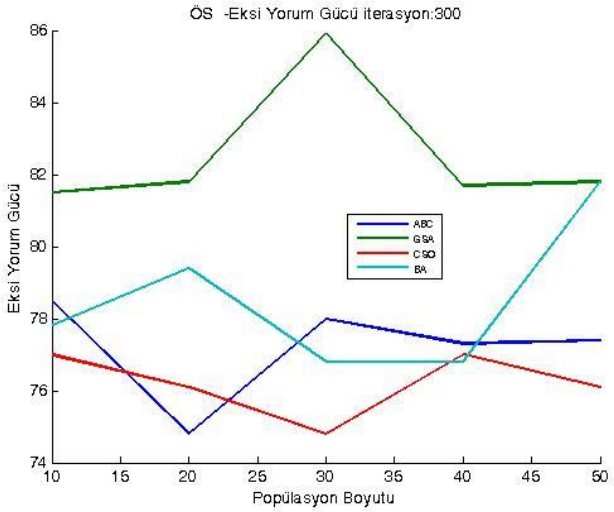
Şekil 1.51. QSAR veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



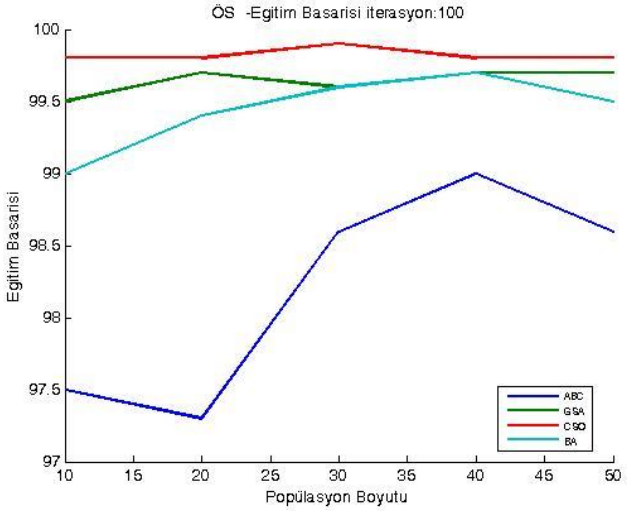
Şekil 1.52. QSAR veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



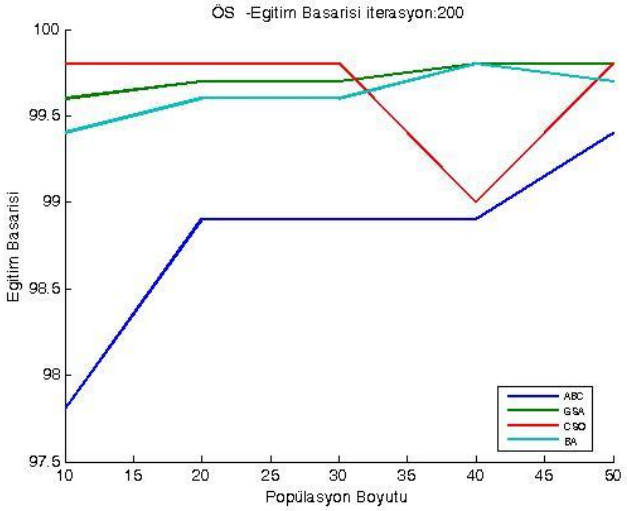
Şekil 1.53. QSAR veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



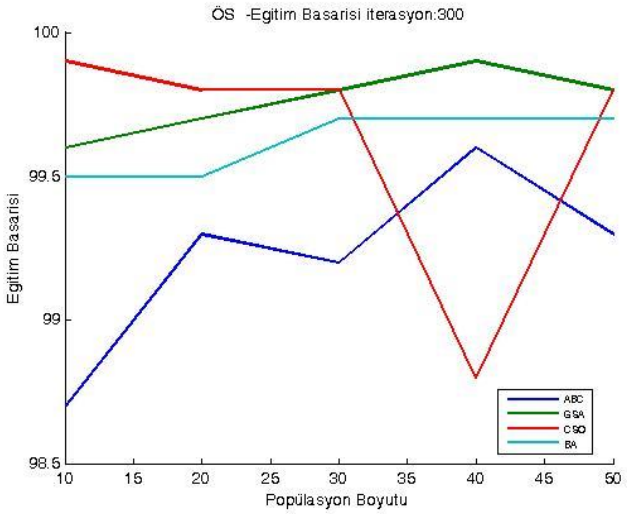
Şekil 1.54. QSAR veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



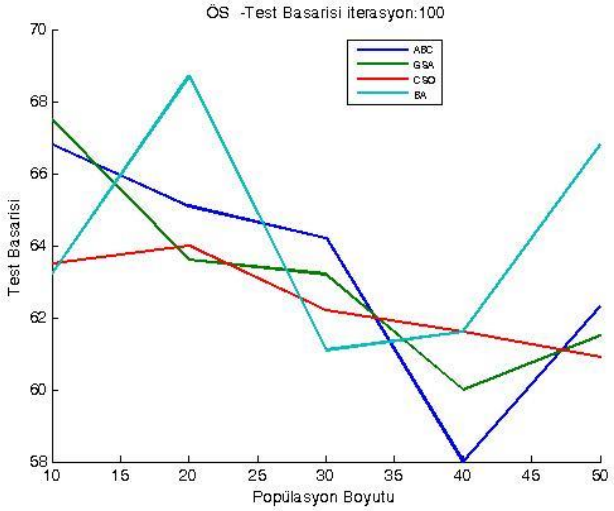
Şekil 1.55. DRD veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



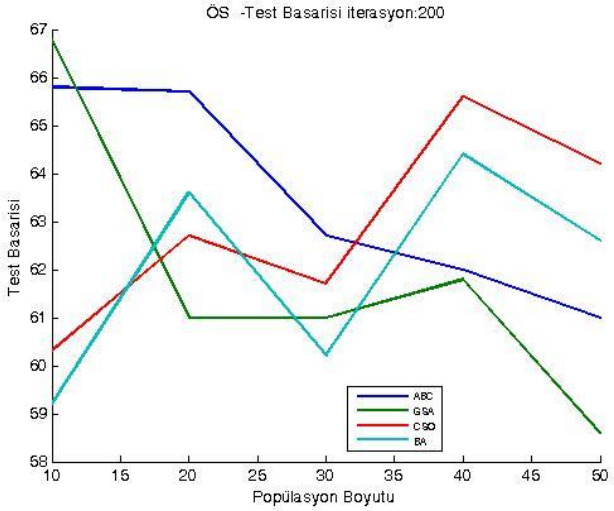
Şekil 1.56. DRD veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



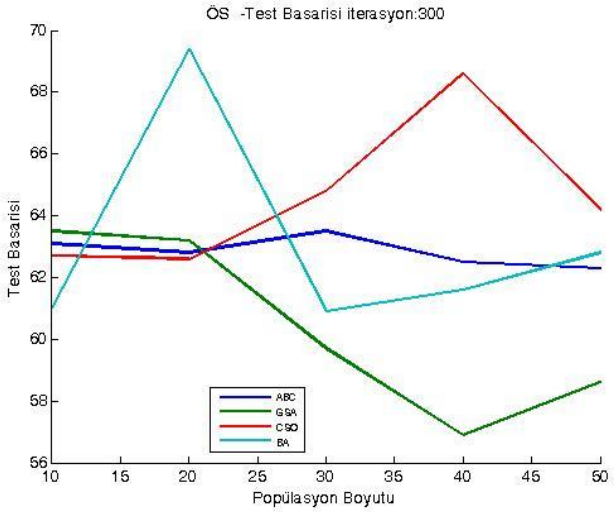
Şekil 1.57. DRD veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



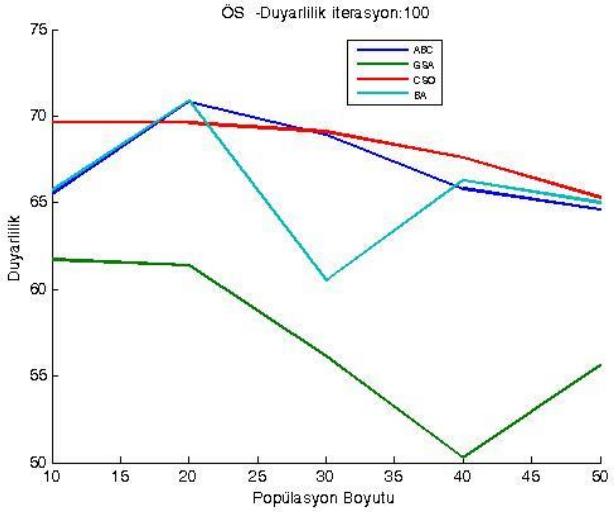
Şekil 1.58. DRD veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



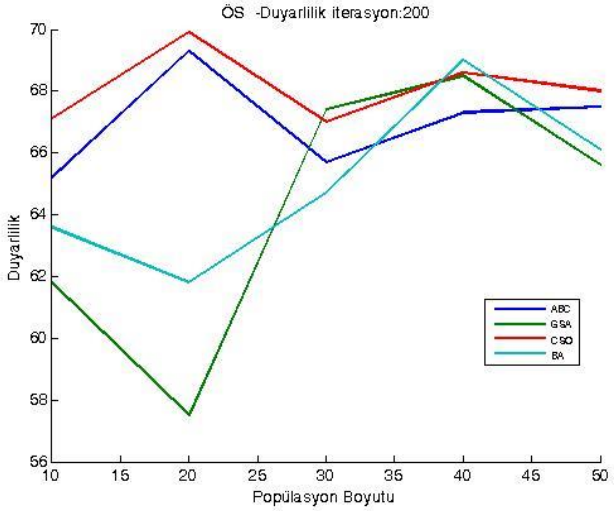
Şekil 1.59. DRD veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



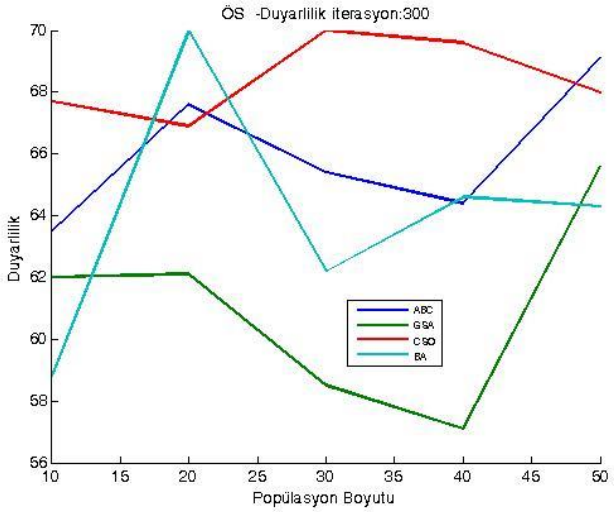
Şekil 1.60. DRD veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



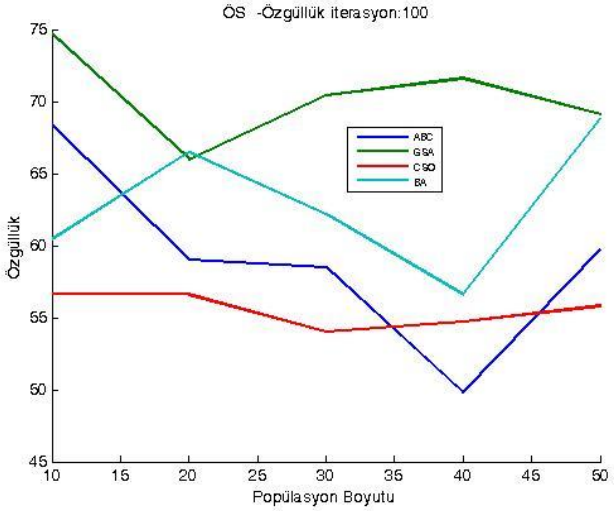
Şekil 1.61. DRD veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



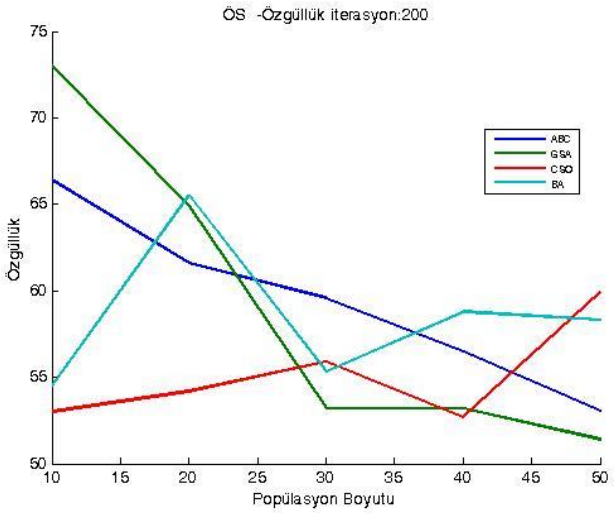
Şekil 1.62. DRD veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



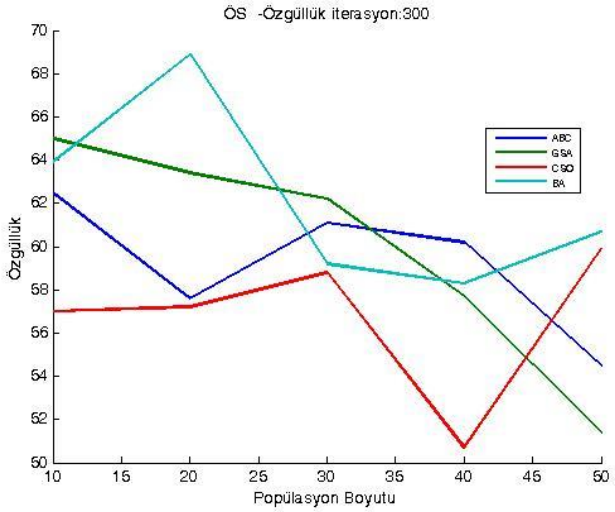
Şekil 1.63. DRD veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



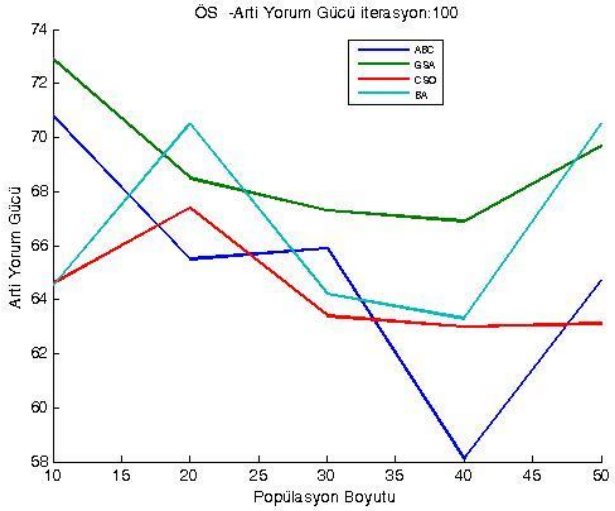
Şekil 1.64. DRD veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



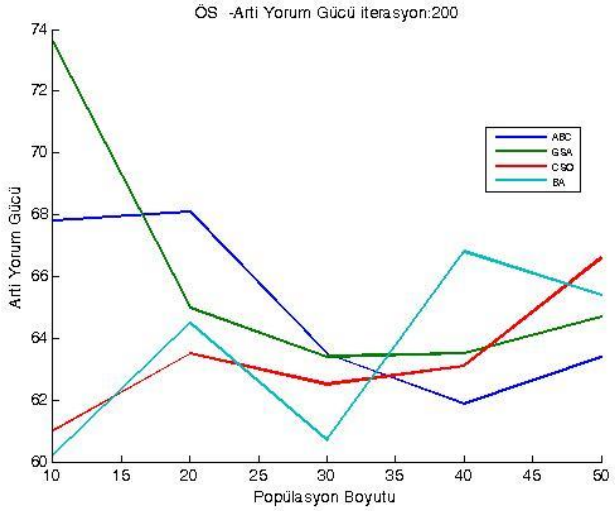
Şekil 1.65. DRD veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



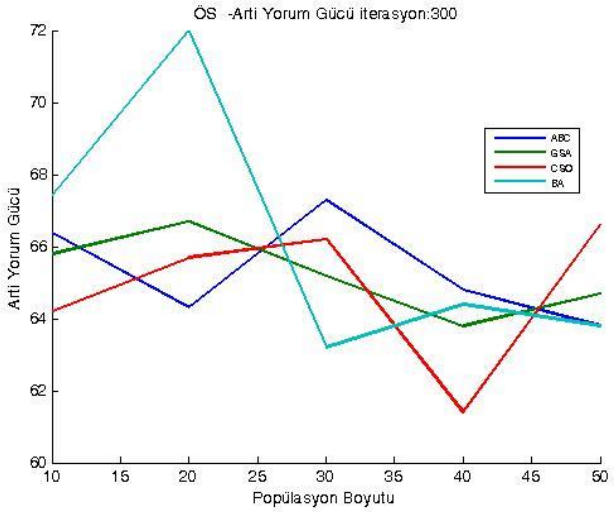
Şekil 1.66. DRD veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



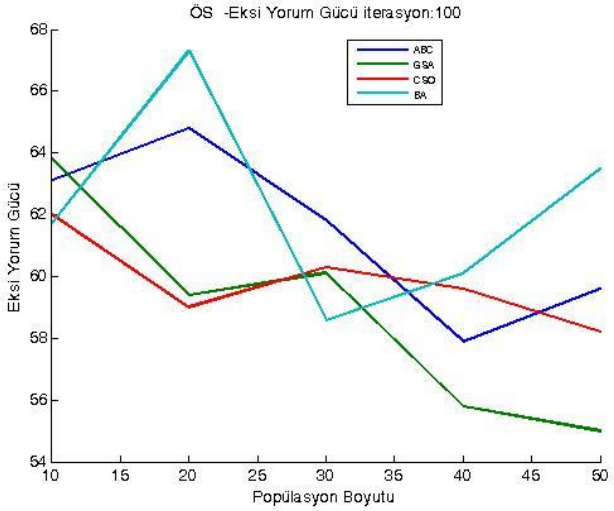
Şekil 1.67. DRD veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



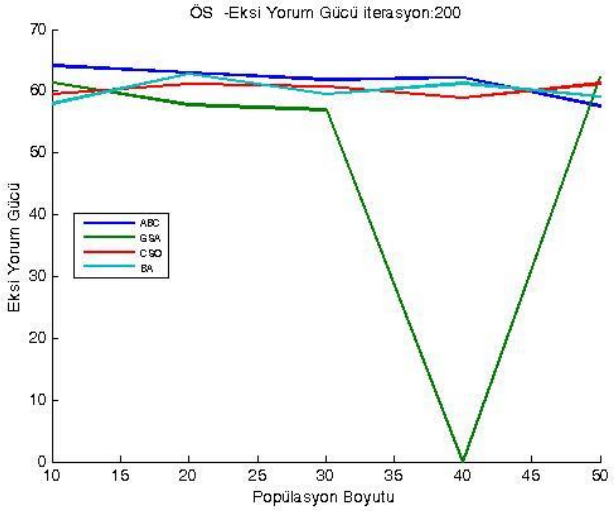
Şekil 1.68. DRD veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



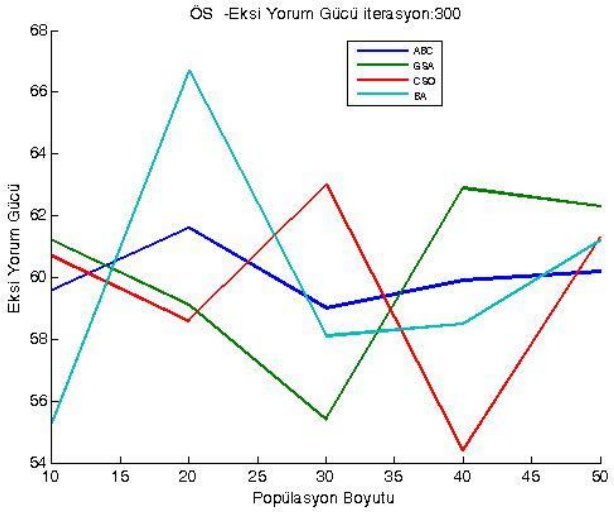
Şekil 1.69. DRD veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



Şekil 1.70. DRD veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları

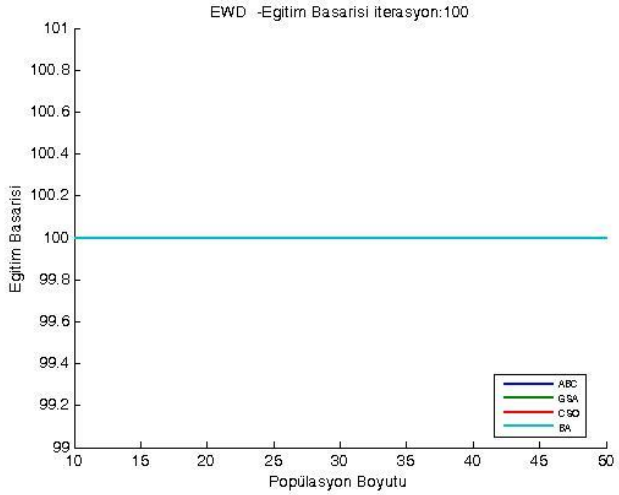


Şekil 1.71. DRD veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları

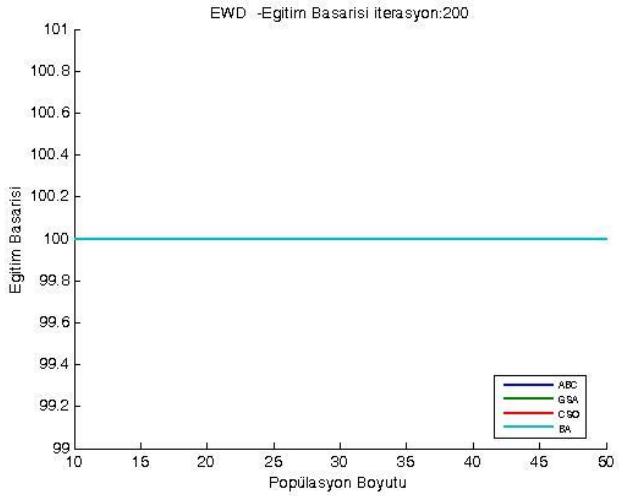


Şekil 1.72. DRD veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları

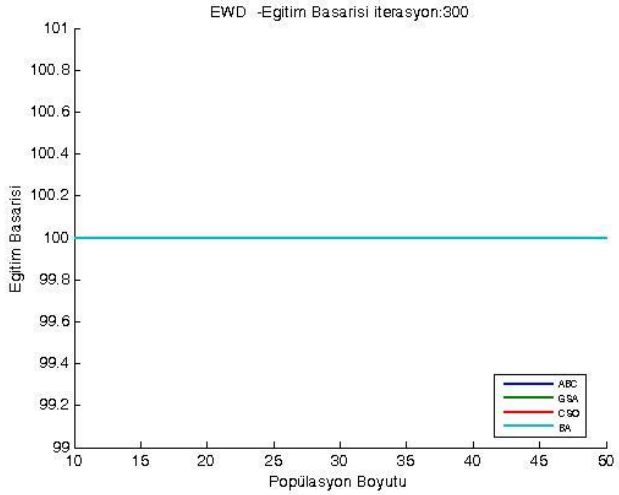
EK -2 EWD Yöntemi İçin Test Sonuçları



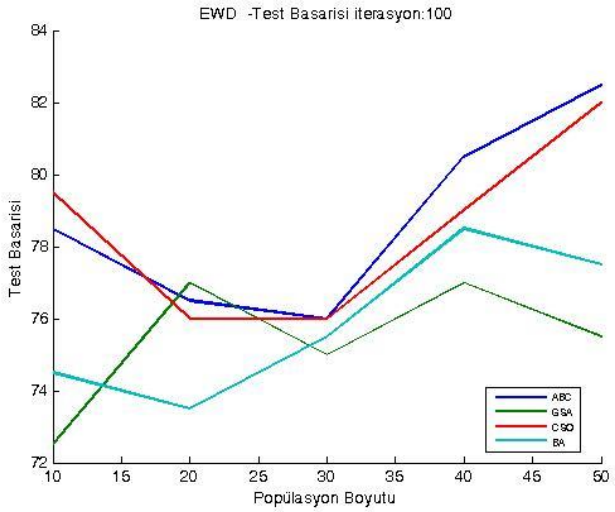
Şekil 2.1. WPBC veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



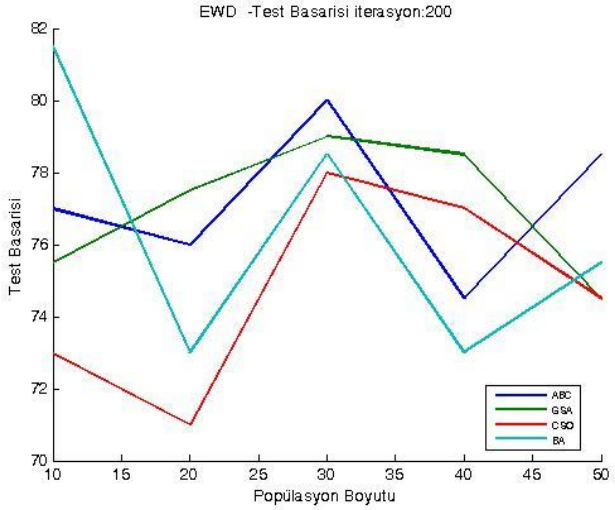
Şekil 2.2. WPBC veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



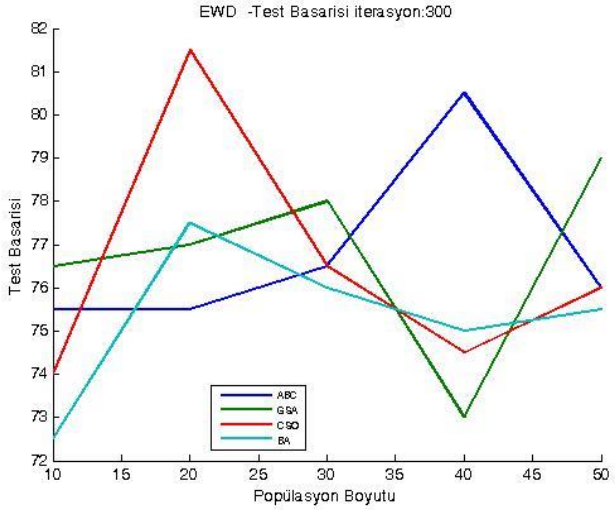
Şekil 2.3. WPBC veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



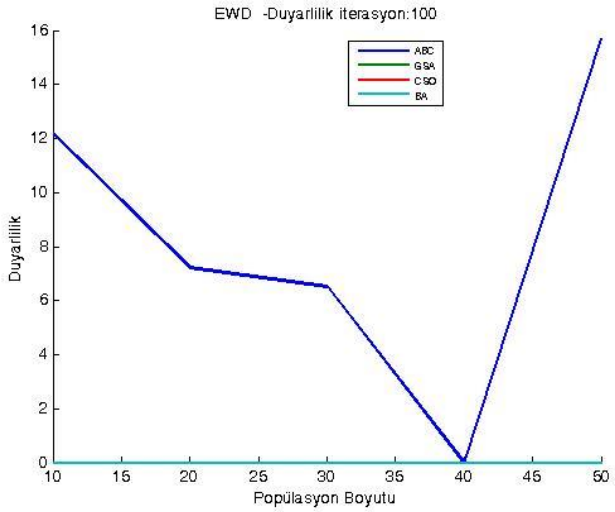
Şekil 2.4. WPBC veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



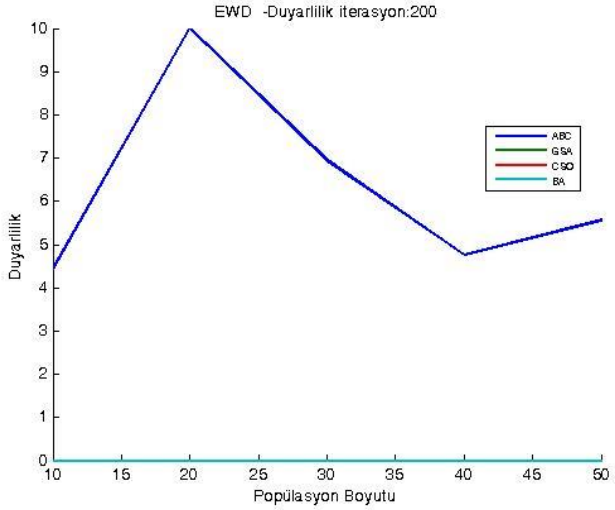
Şekil 2.5. WPBC veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



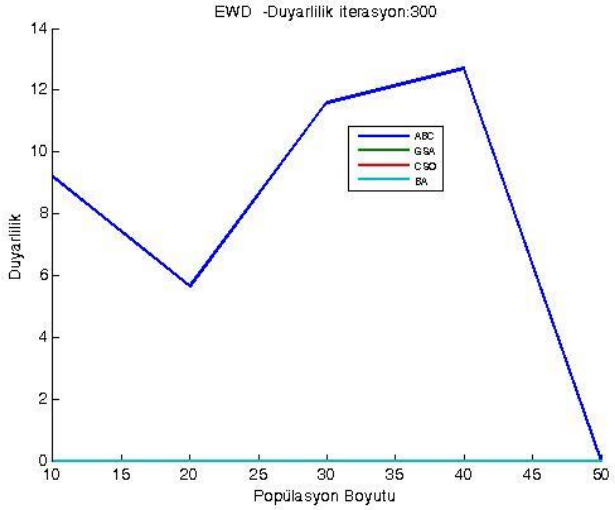
Şekil 2.6. WPBC veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



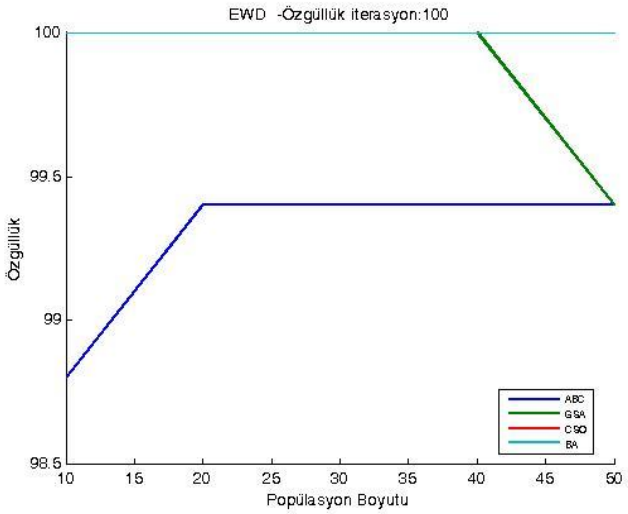
Şekil 2.7. WPBC veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



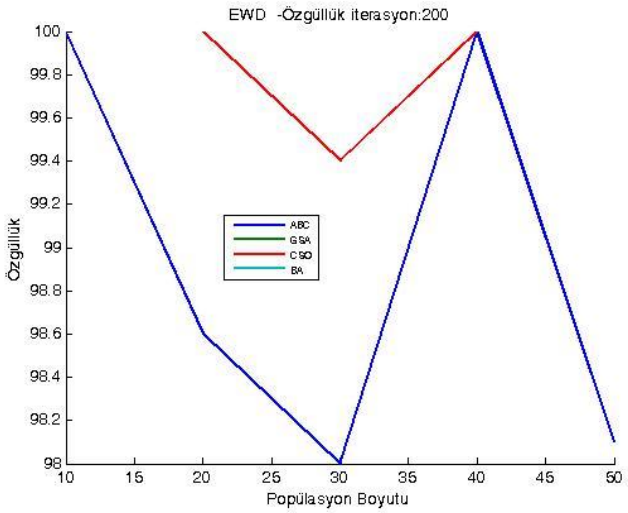
Şekil 2.8. WPBC veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



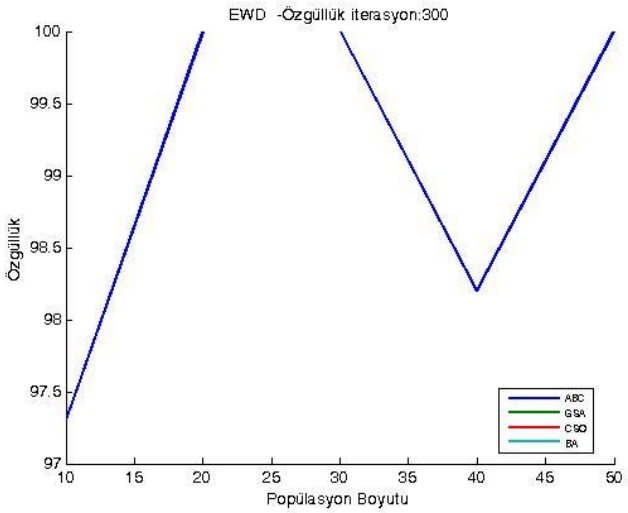
Şekil 2.9. WPBC veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



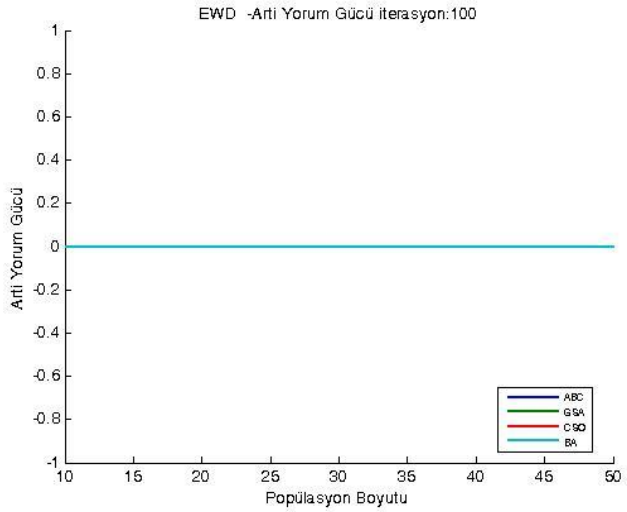
Şekil 2.10. WPBC veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



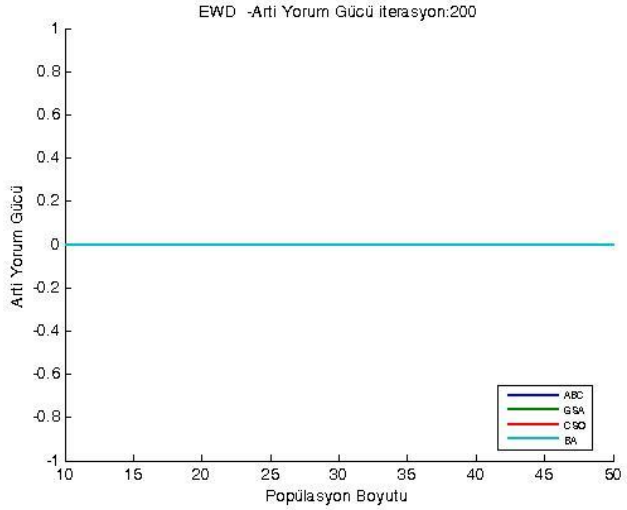
Şekil 2.11. WPBC veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



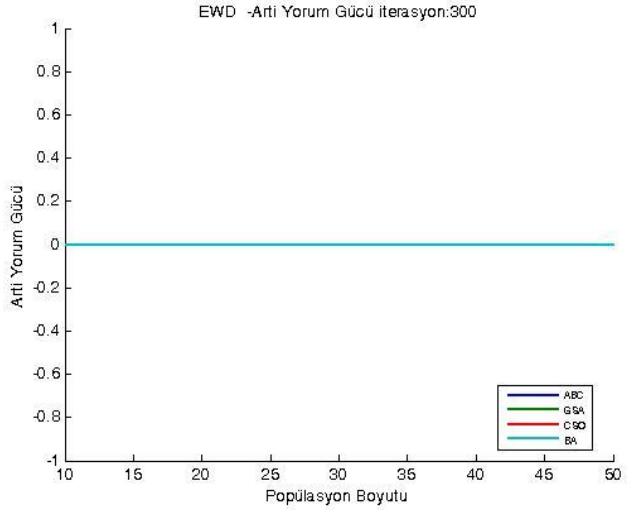
Şekil 2.12. WPBC veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



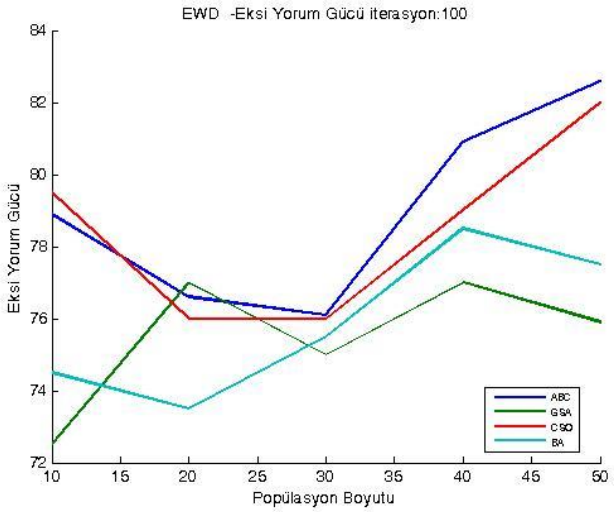
Şekil 2.13. WPBC veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



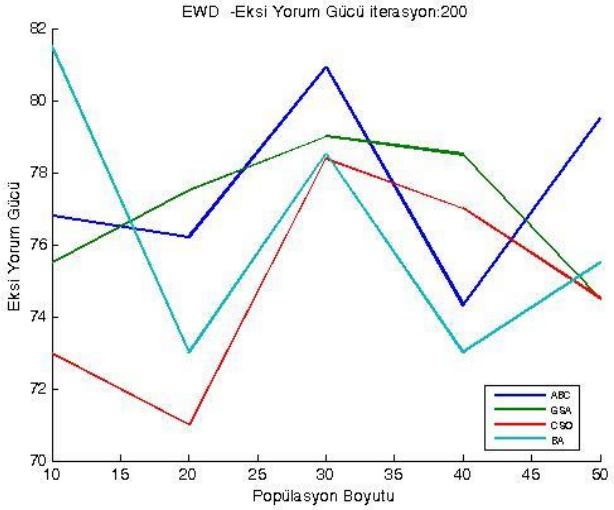
Şekil 2.14. WPBC veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



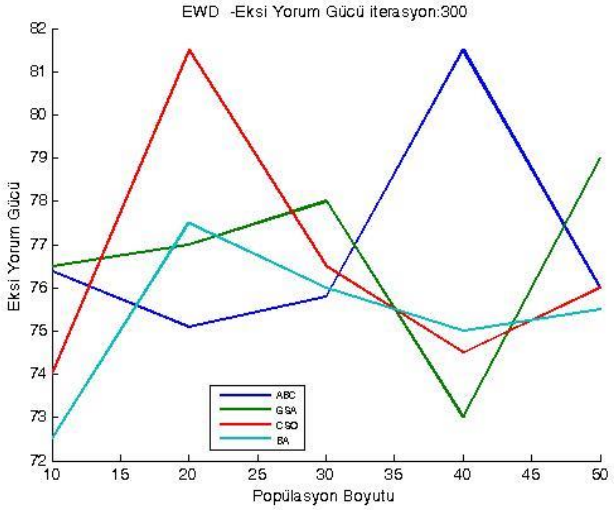
Şekil 2.15. WPBC veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



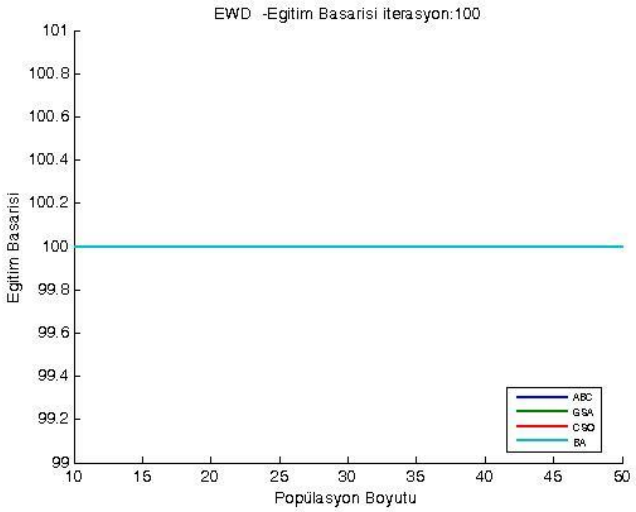
Şekil 2.16. WPBC veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



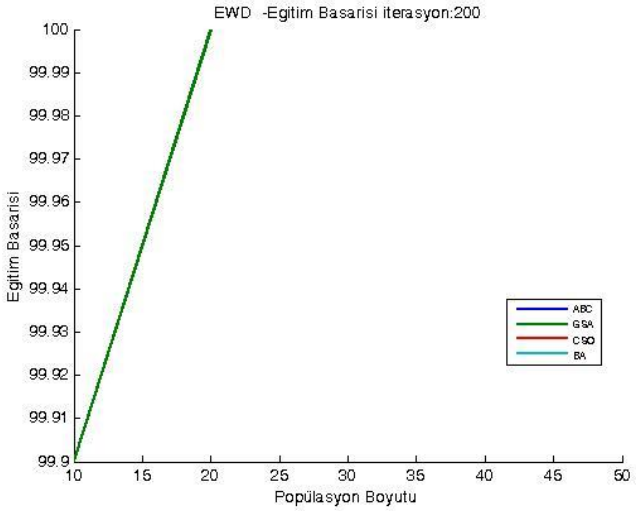
Şekil 2.17. WPBC veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



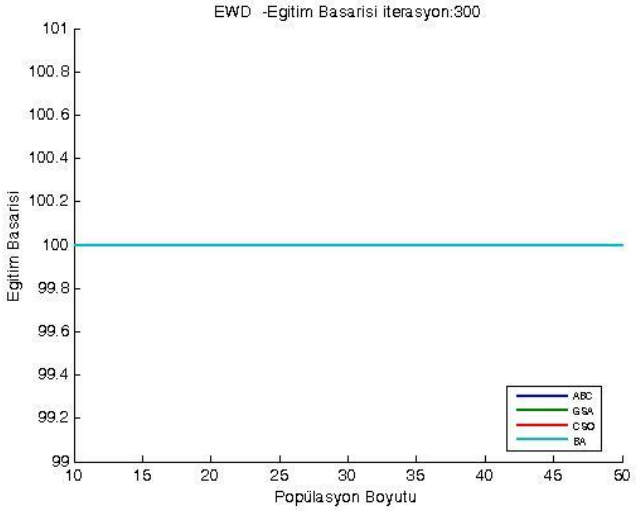
Şekil 2.18. WPBC veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



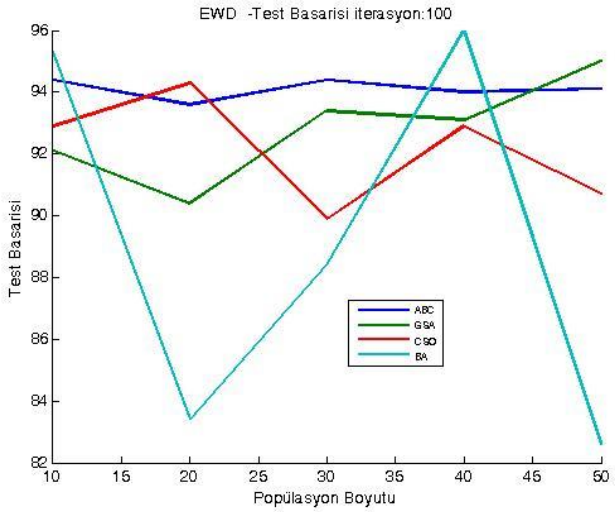
Şekil 2.19. BCW veri kümesi üzerinde 100 iterasyon için ortalama eğitim sonuçları



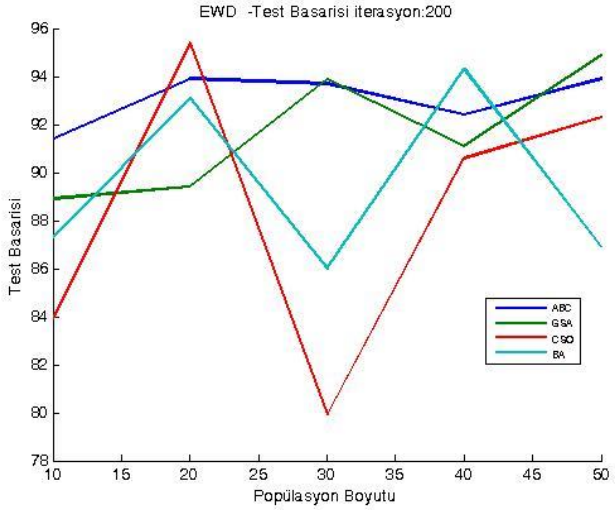
Şekil 2.20. BCW veri kümesi üzerinde 200 iterasyon için ortalama eğitim sonuçları



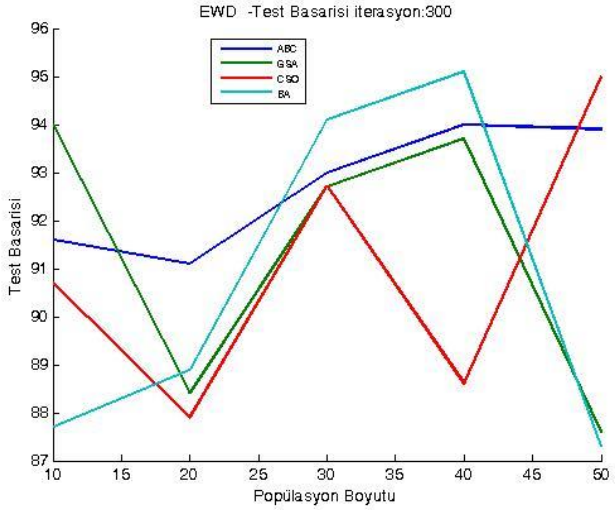
Şekil 2.21. BCW veri kümesi üzerinde 300 iterasyon için ortalama eğitim sonuçları



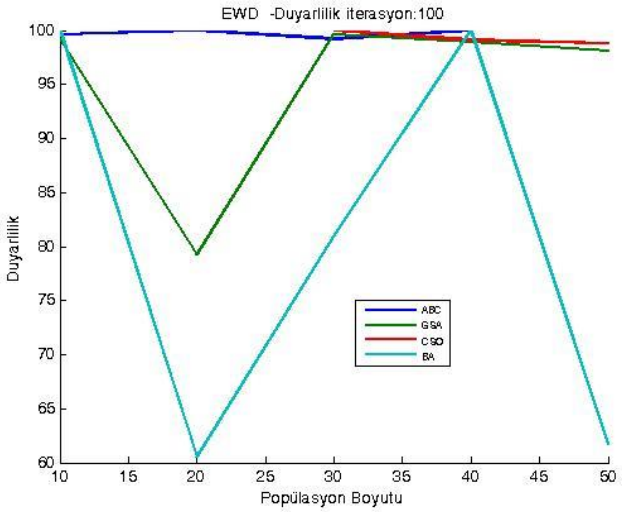
Şekil 2.22. BCW veri kümesi üzerinde 100 iterasyon için ortalama test sonuçları



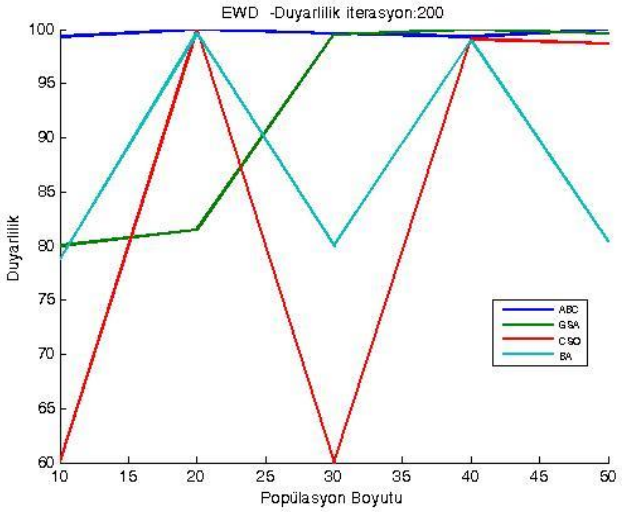
Şekil 2.23. BCW veri kümesi üzerinde 200 iterasyon için ortalama test sonuçları



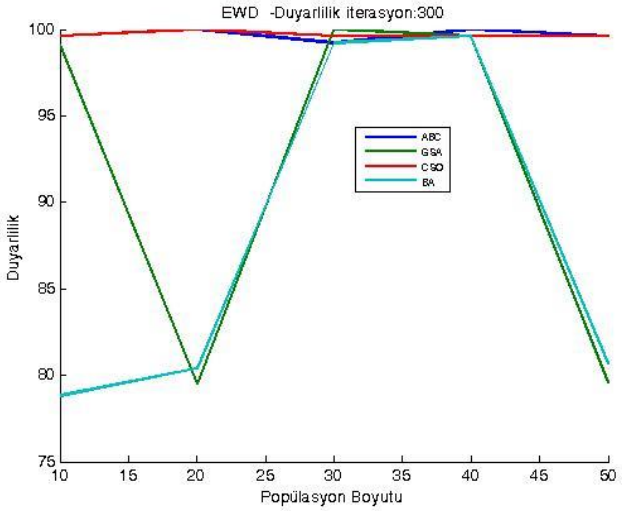
Şekil 2.24. BCW veri kümesi üzerinde 300 iterasyon için ortalama test sonuçları



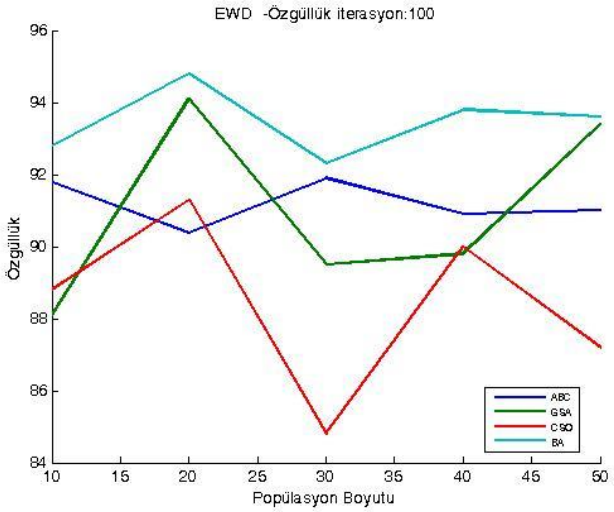
Şekil 2.25. BCW veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



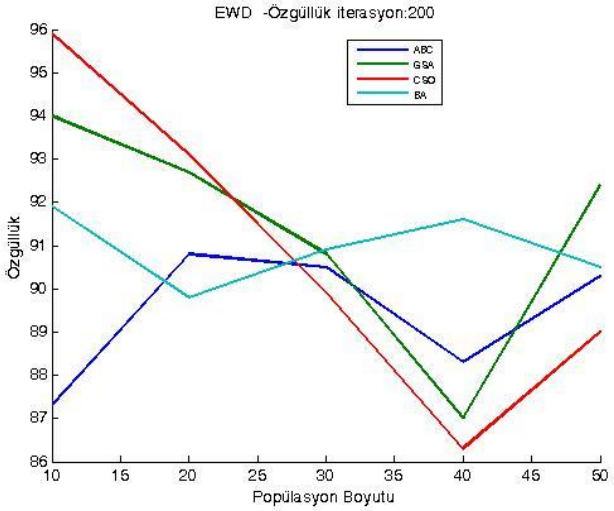
Şekil 2.26. BCW veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



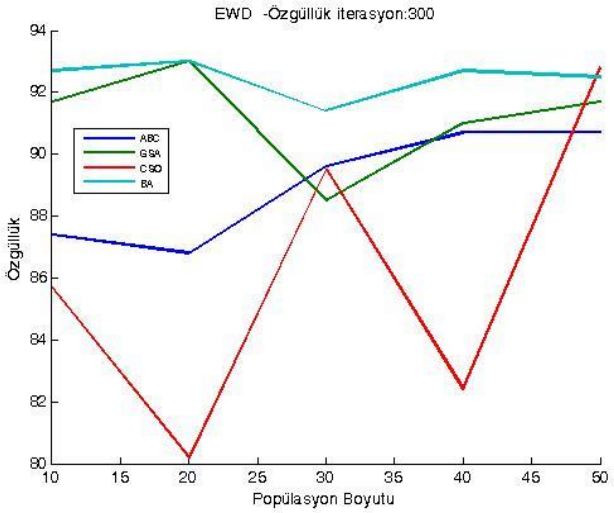
Şekil 2.27. BCW veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



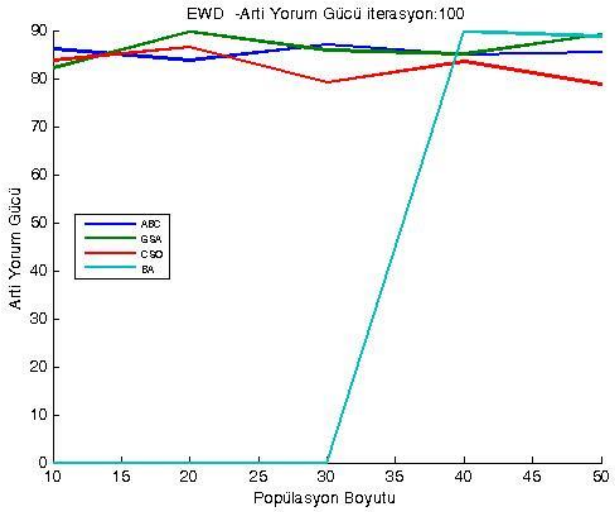
Şekil 2.28. BCW veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



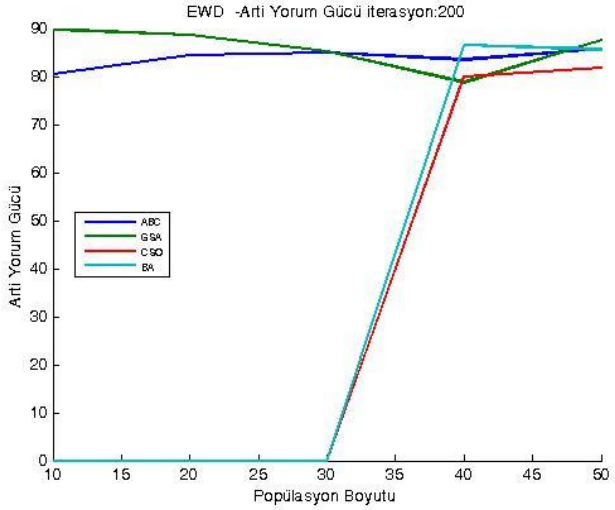
Şekil 2.29. BCW veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



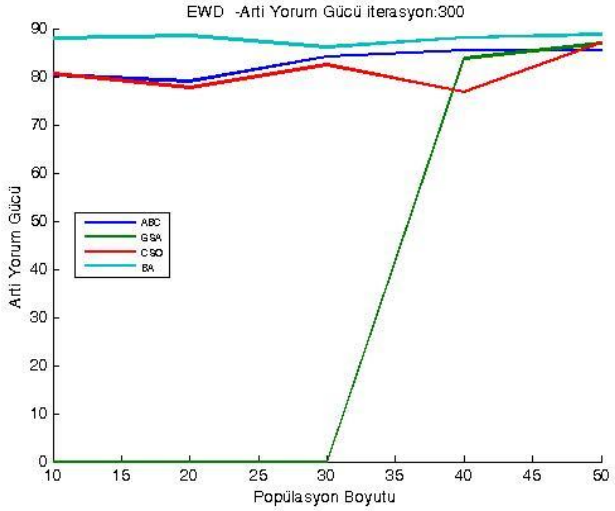
Şekil 2.30. BCW veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



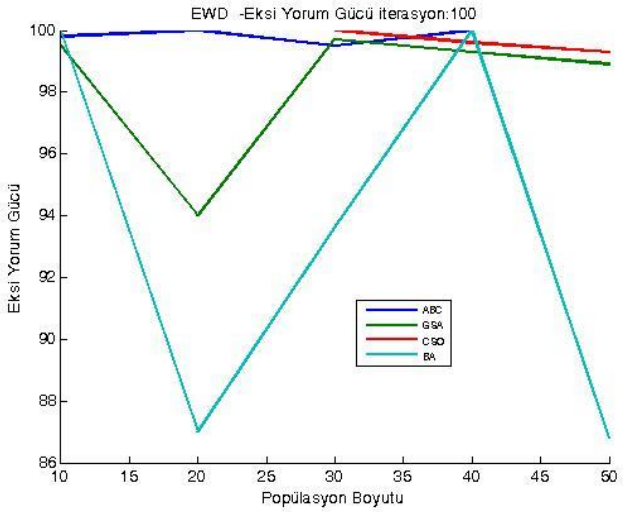
Şekil 2.31. BCW veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



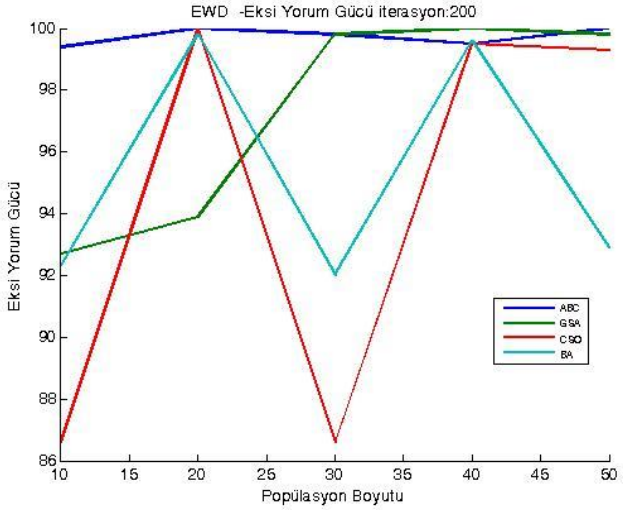
Şekil 2.32. BCW veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



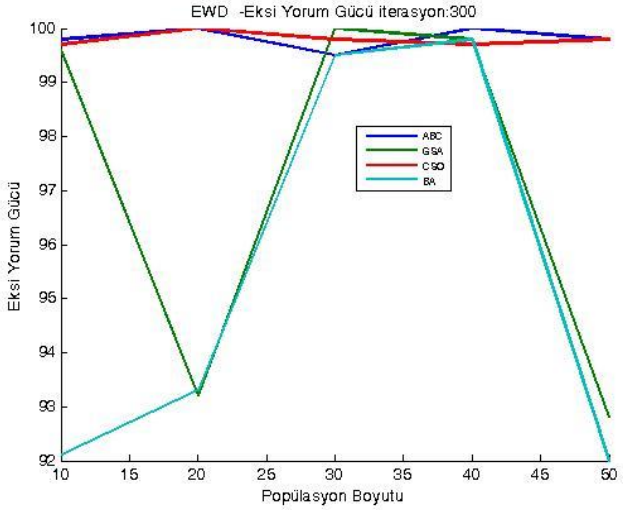
Şekil 2.33. BCW veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



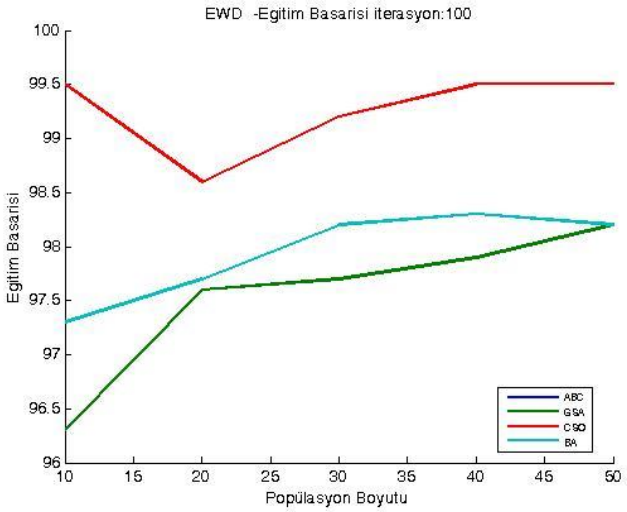
Şekil 2.34. BCW veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



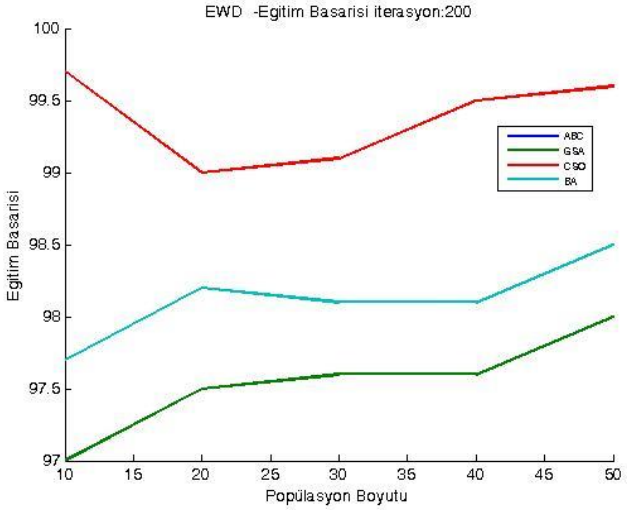
Şekil 2.35. BCW veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



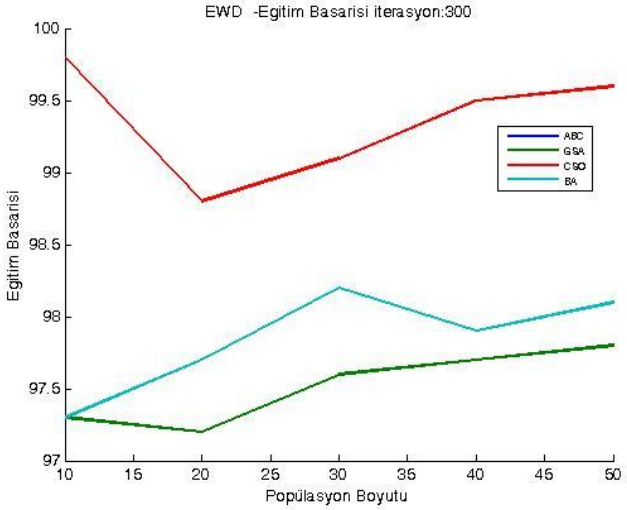
Şekil 2.36. BCW veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



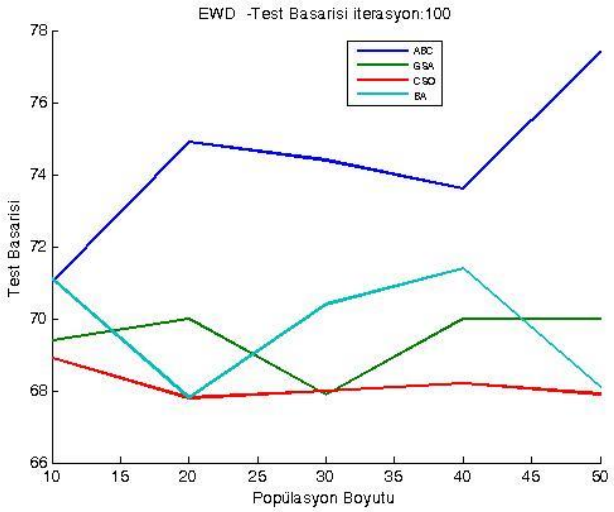
Şekil 2.37. QSAR veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



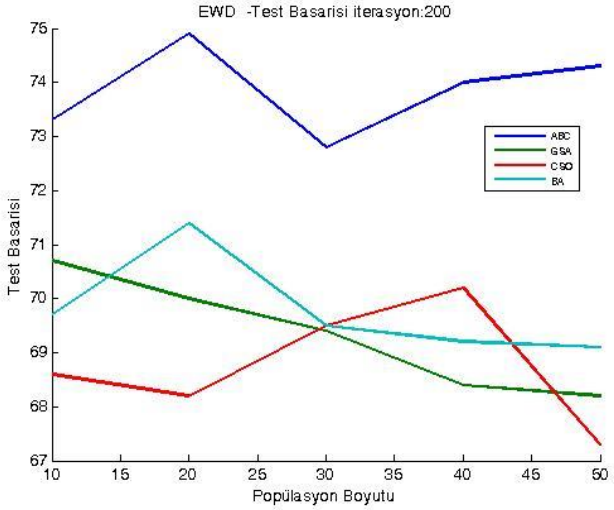
Şekil 2.38. QSAR veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



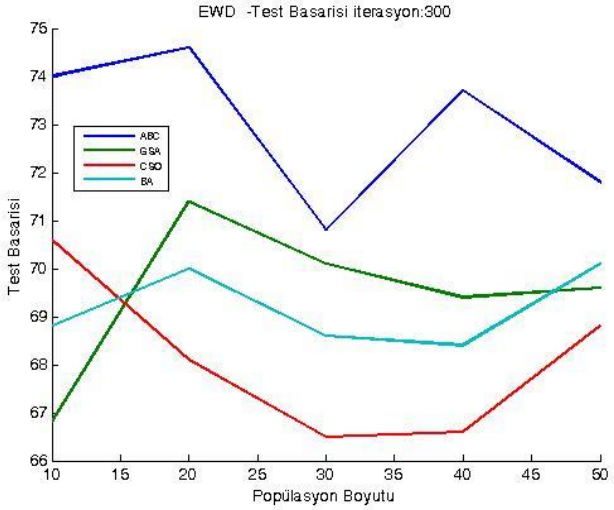
Şekil 2.39. QSAR veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



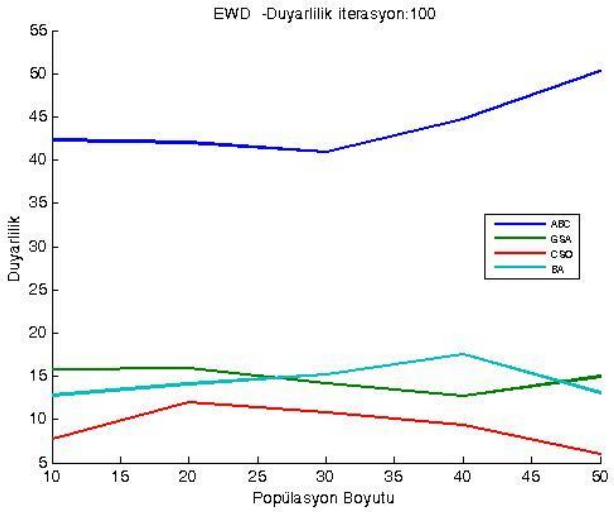
Şekil 2.40. QSAR veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



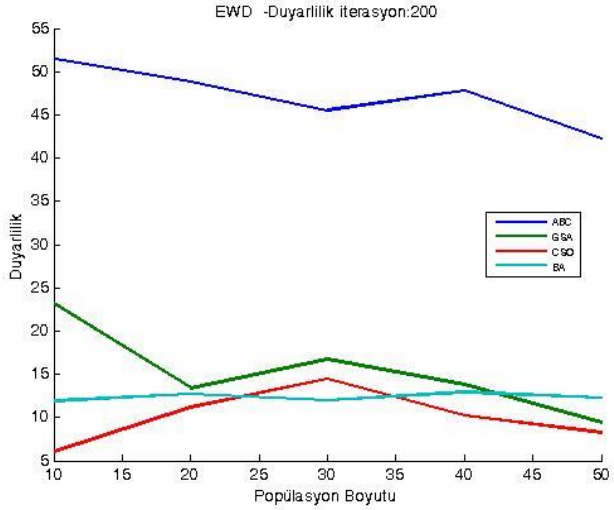
Şekil 2.41. QSAR veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



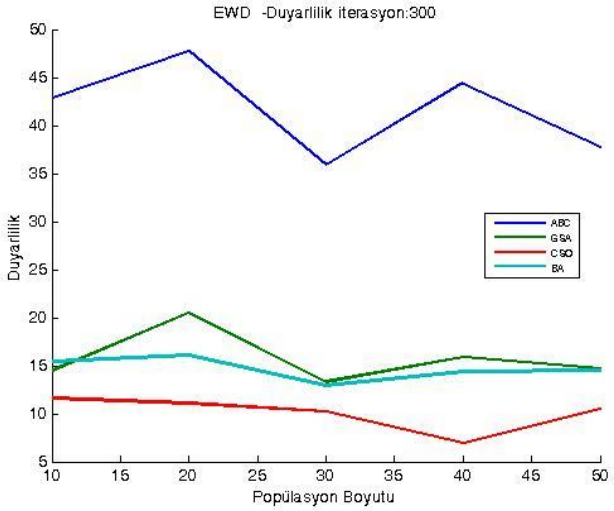
Şekil 2.42. QSAR veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



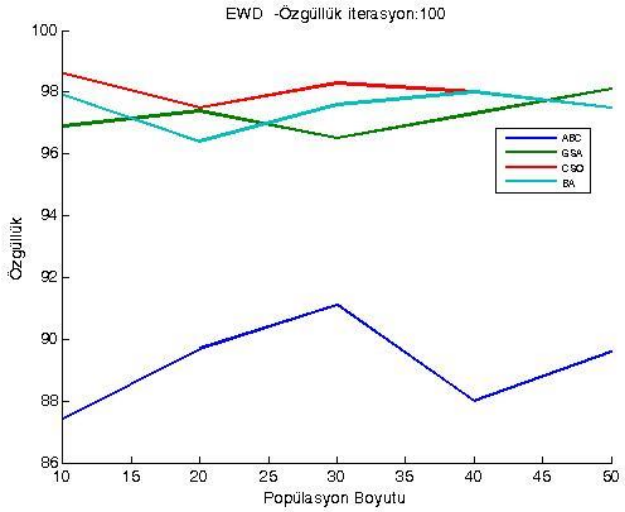
Şekil 2.43. QSAR veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



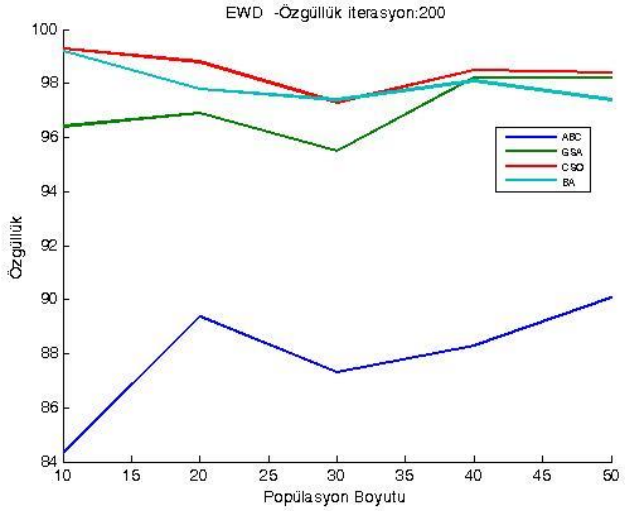
Şekil 2.44. QSAR veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



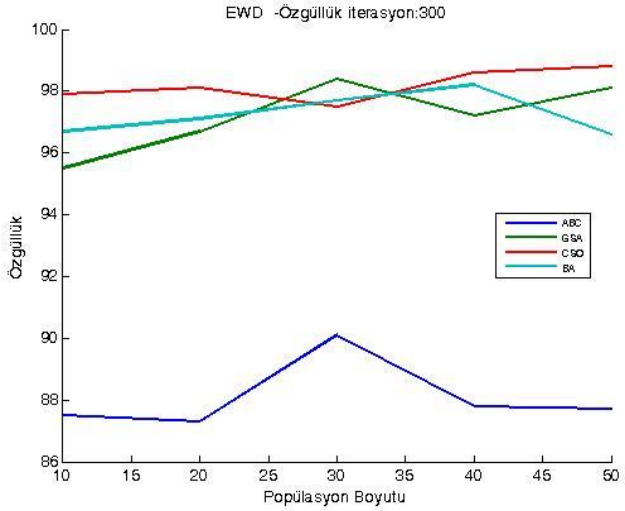
Şekil 2.45. QSAR veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



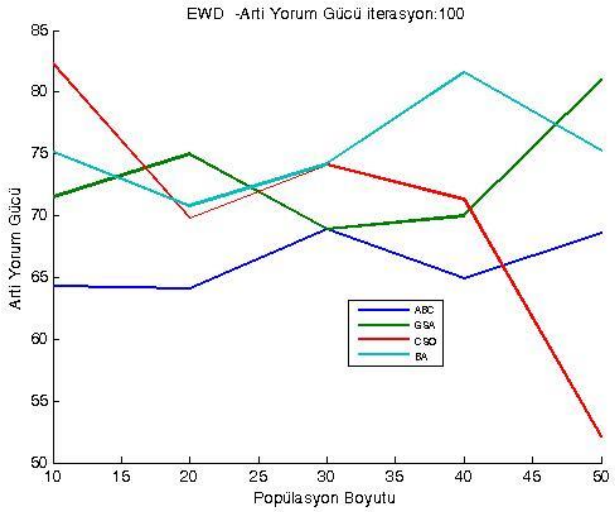
Şekil 2.46. QSAR veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



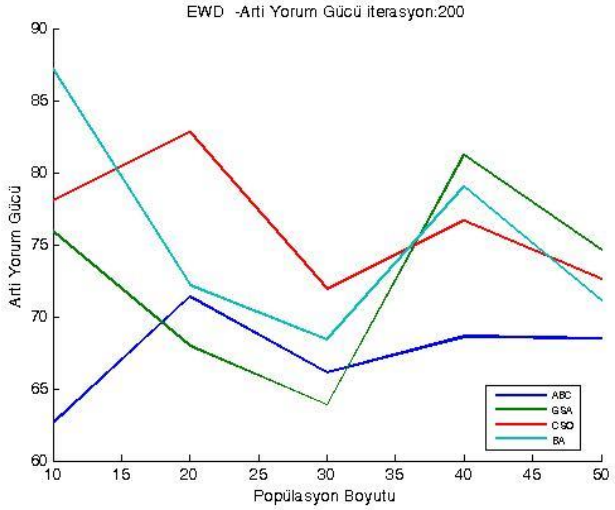
Şekil 2.47. QSAR veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



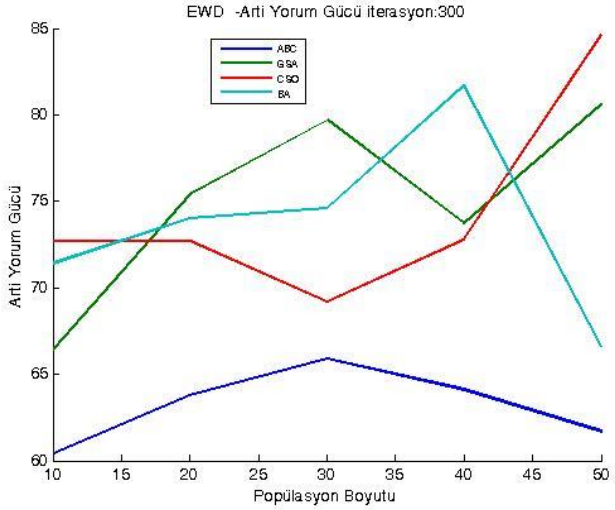
Şekil 2.48. QSAR veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



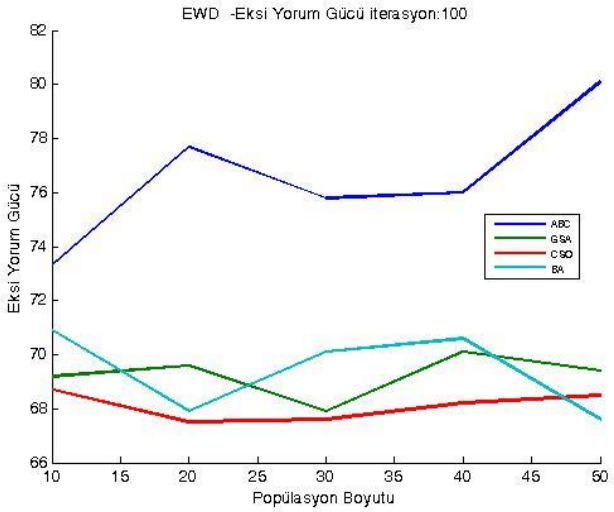
Şekil 2.49. QSAR veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



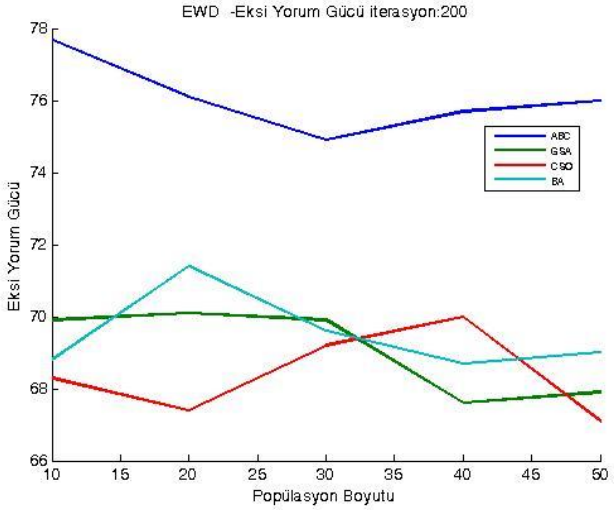
Şekil 2.50. QSAR veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



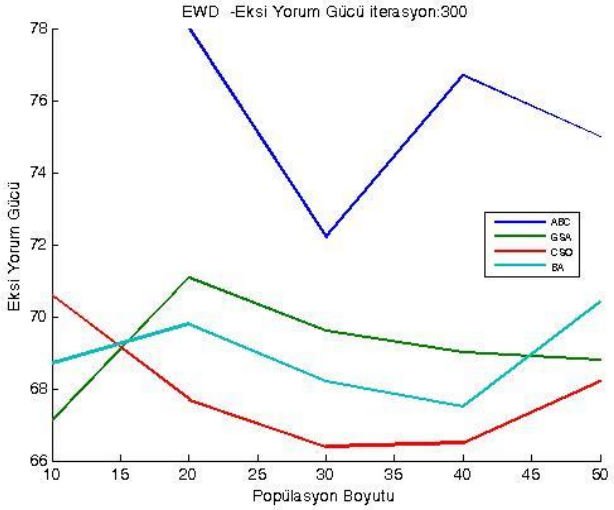
Şekil 2.51. QSAR veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



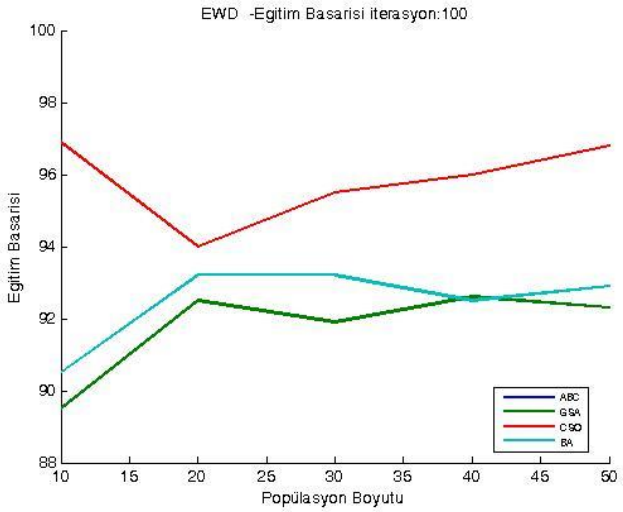
Şekil 2.52. QSAR veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



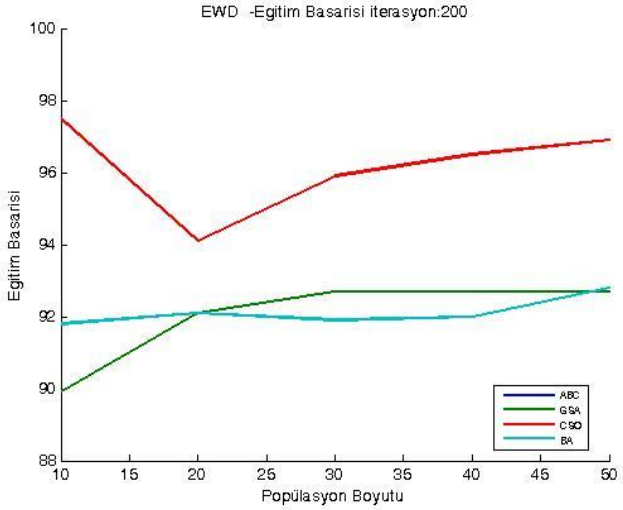
Şekil 2.53. QSAR veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



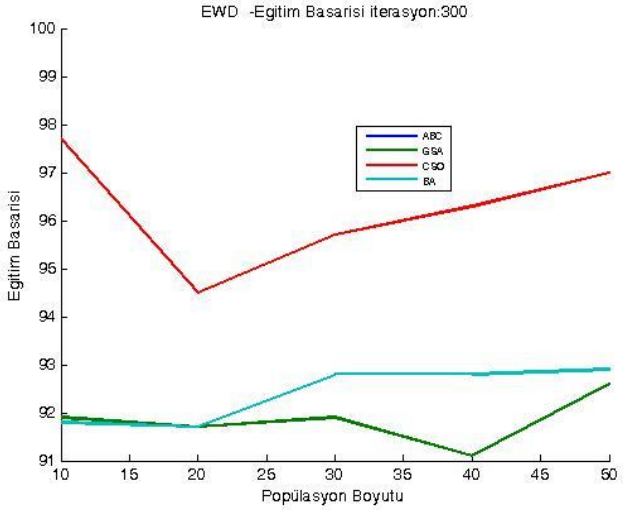
Şekil 2.54. QSAR veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



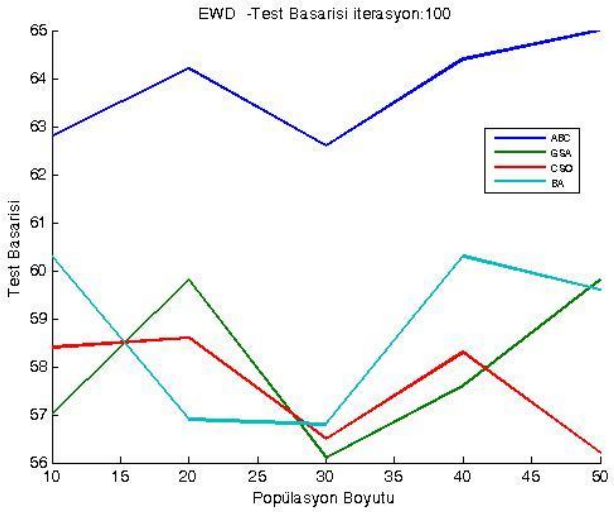
Şekil 2.55. DRD veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



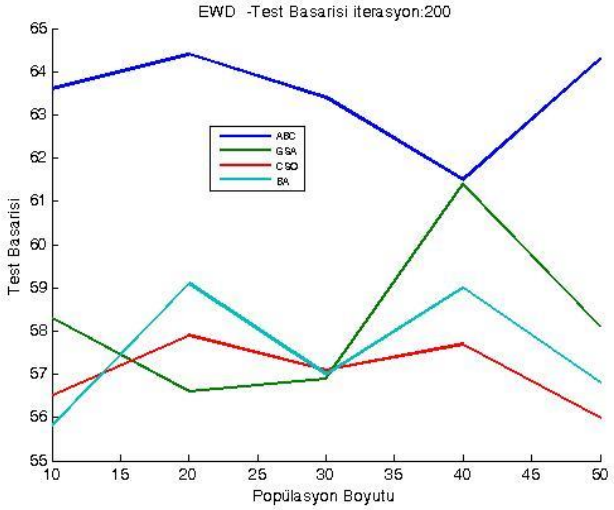
Şekil 2.56. DRD veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



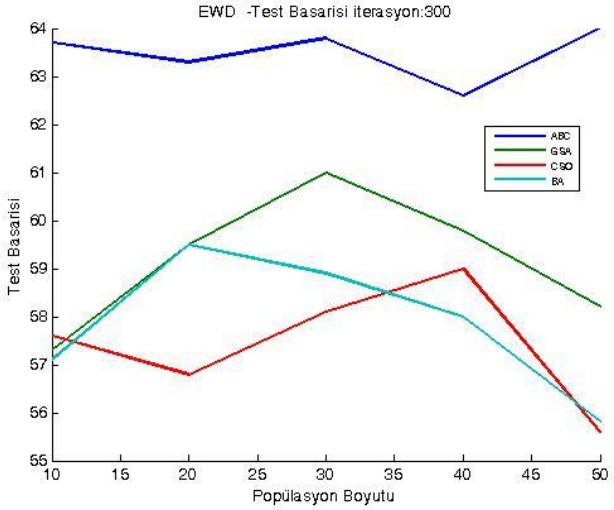
Şekil 2.57. DRD veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



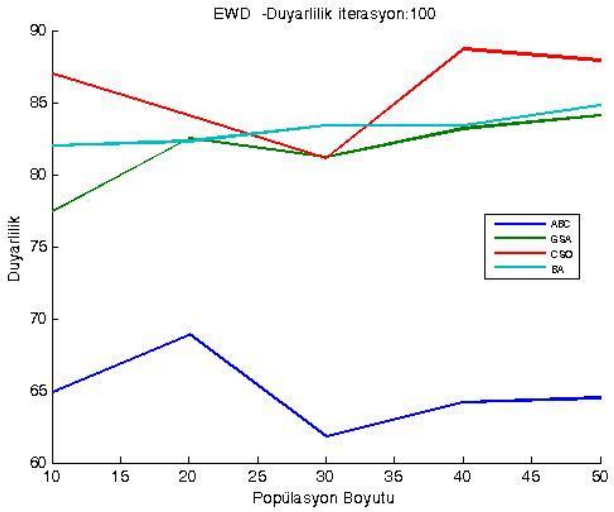
Şekil 2.58. DRD veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



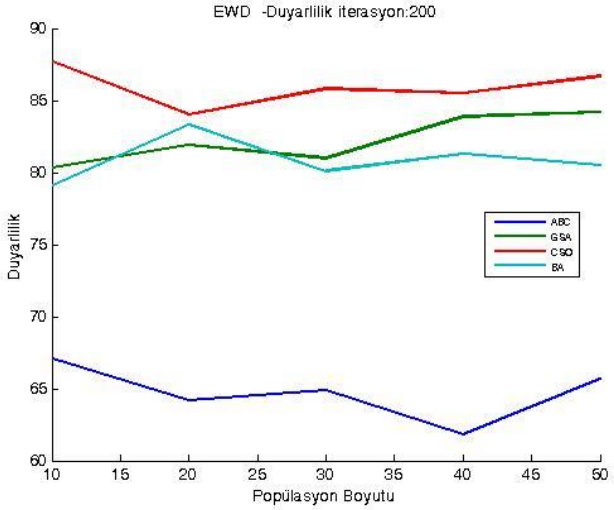
Şekil 2.59. DRD veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



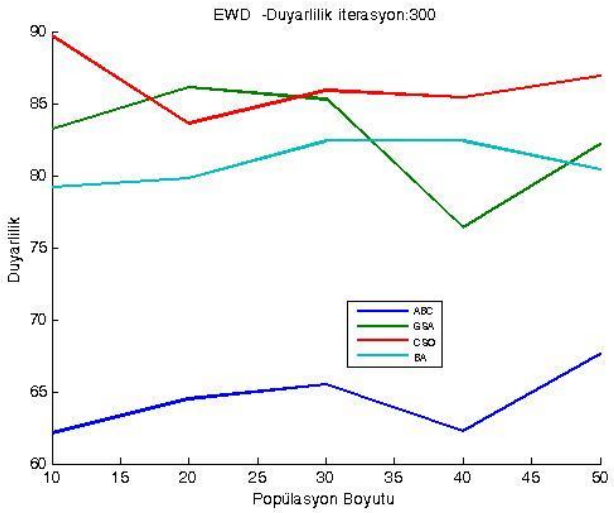
Şekil 2.60. DRD veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



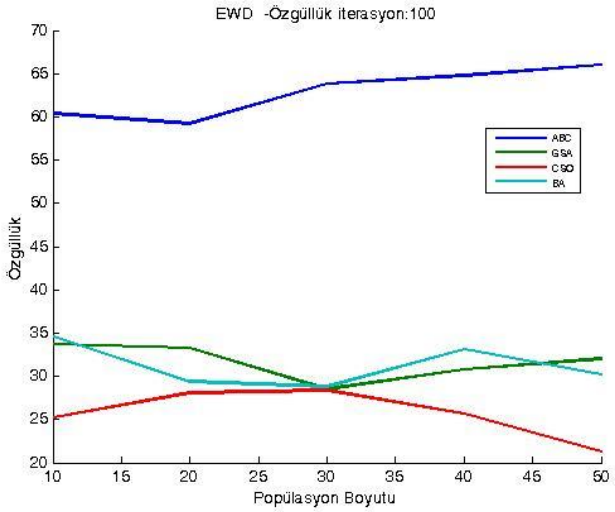
Şekil 2.61. DRD veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



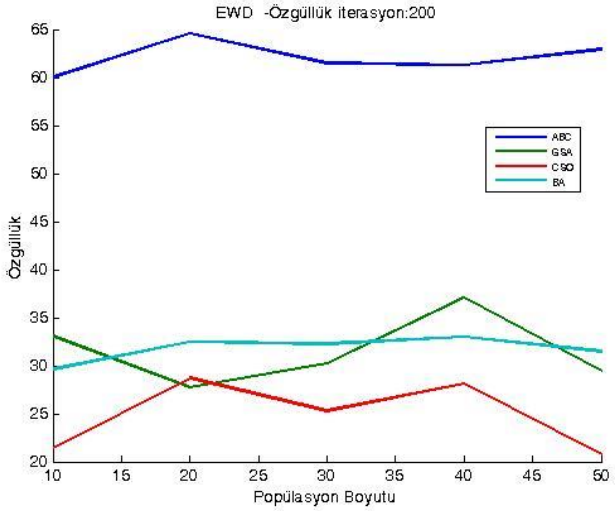
Şekil 2.62. DRD veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



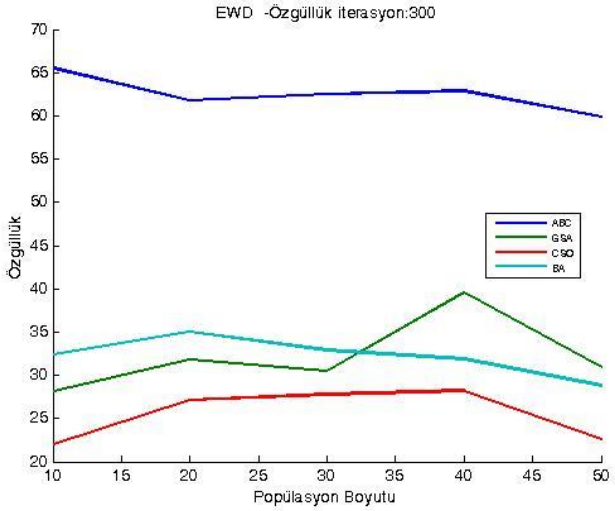
Şekil 2.63. DRD veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



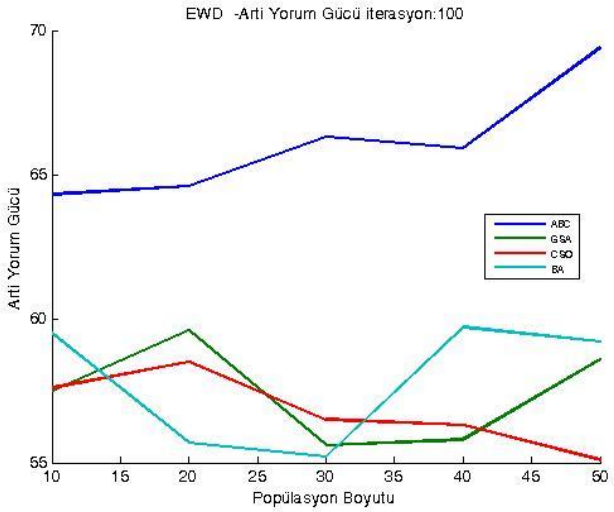
Şekil 2.64. DRD veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



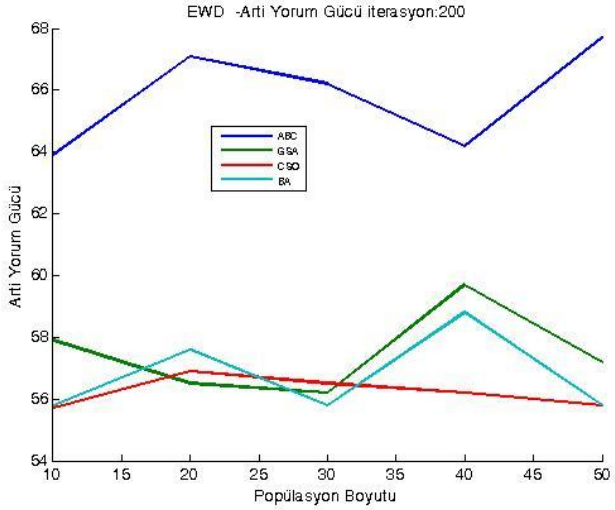
Şekil 2.65. DRD veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



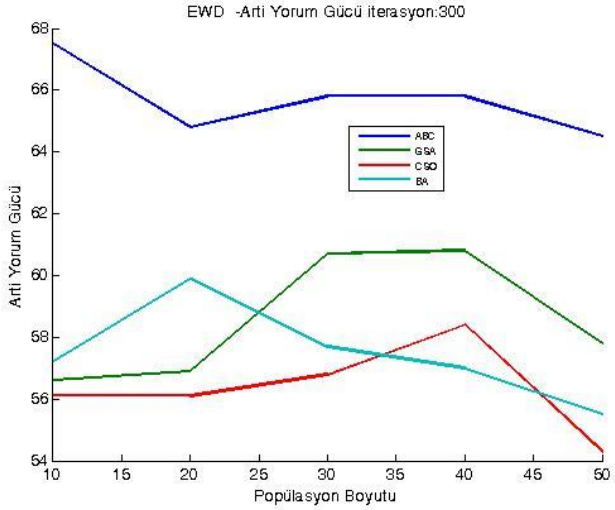
Şekil 2.66. DRD veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



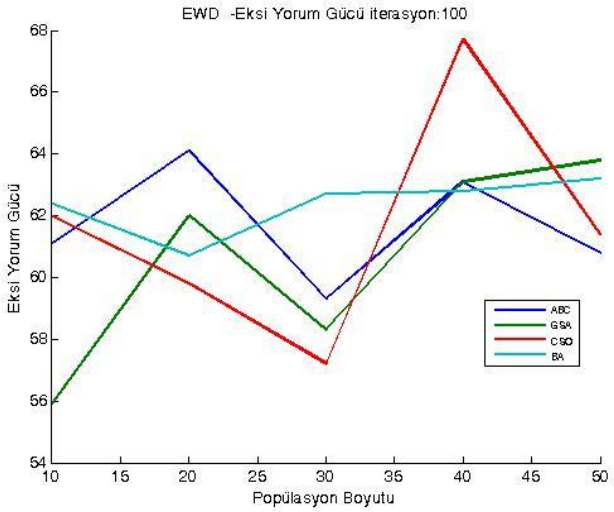
Şekil 2.67. DRD veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



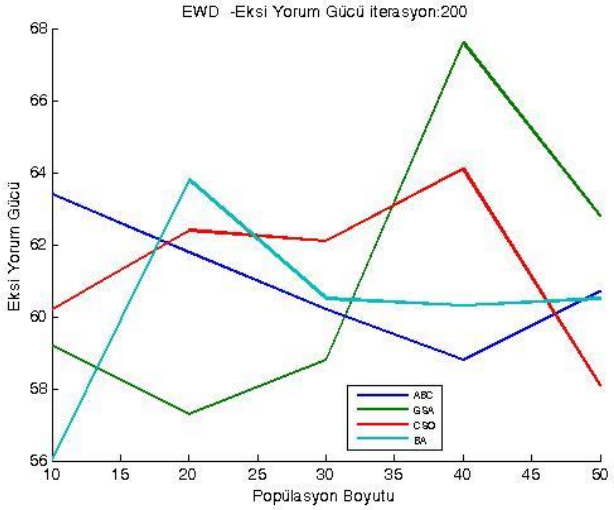
Şekil 2.68. DRD veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



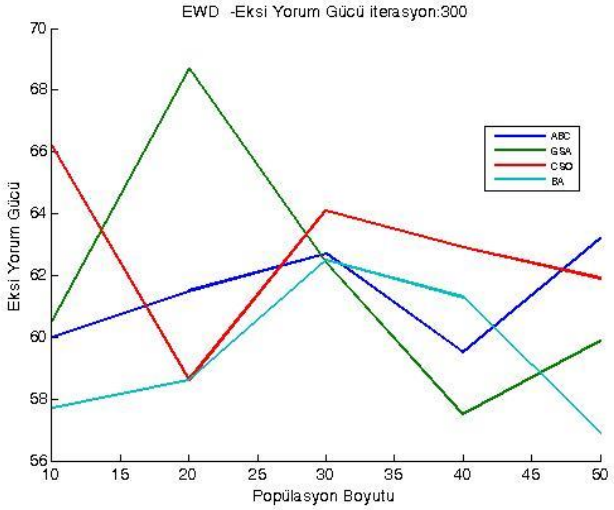
Şekil 2.69. DRD veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



Şekil 2.70. DRD veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları

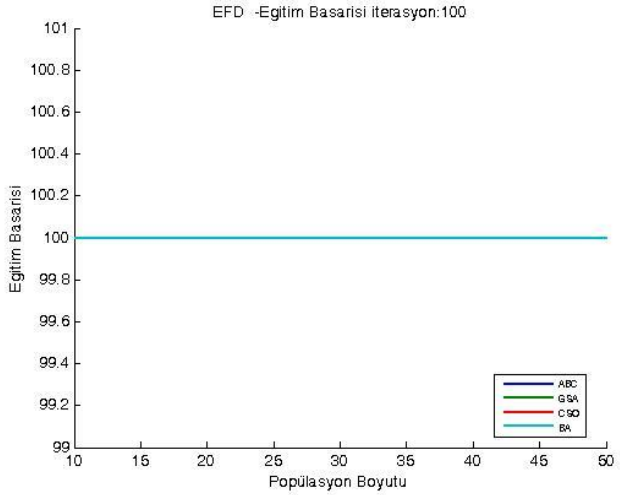


Şekil 2.71. DRD veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları

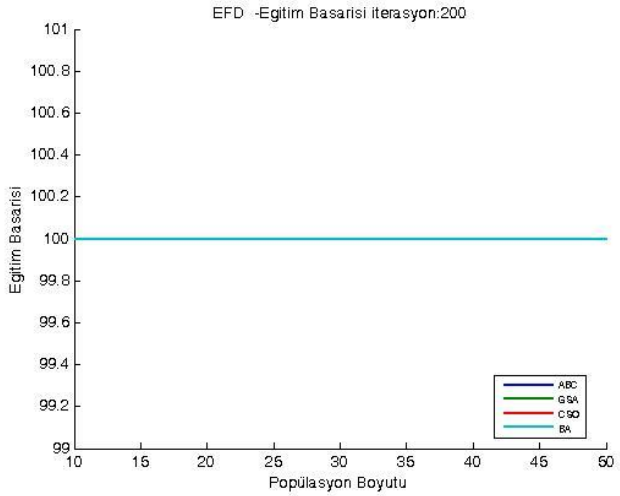


Şekil 2.72. DRD veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları

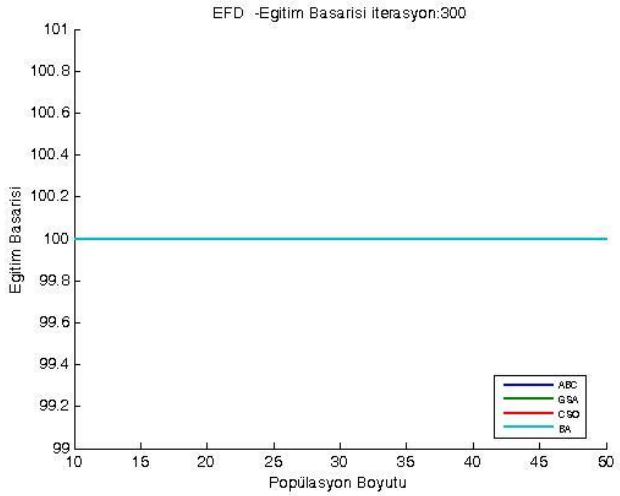
EK - 3 EFD Yöntemi İçin Test Sonuçları



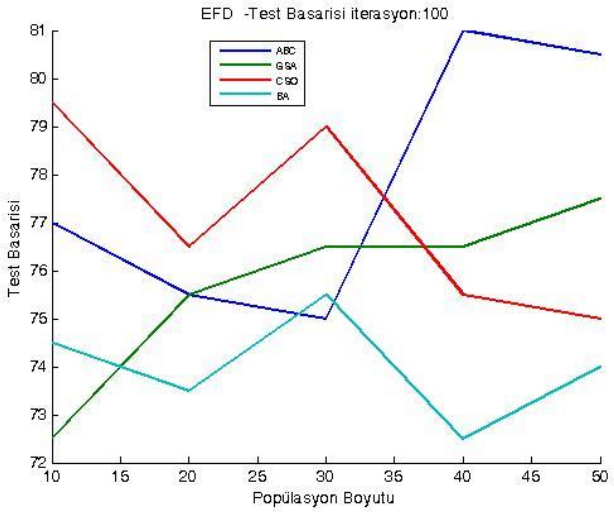
Şekil 3.1. WPBC veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



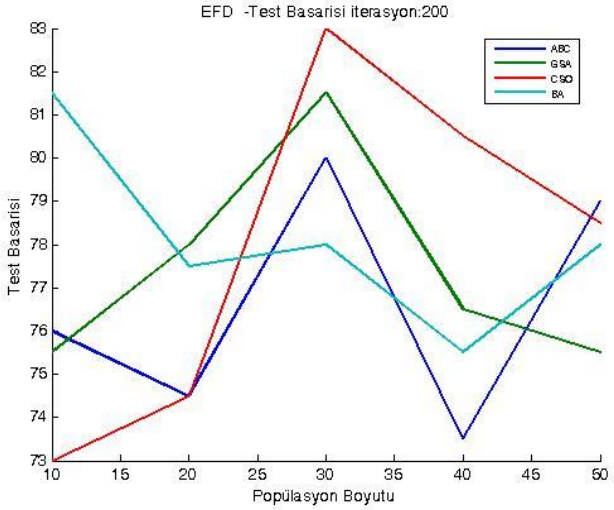
Şekil 3.2. WPBC veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



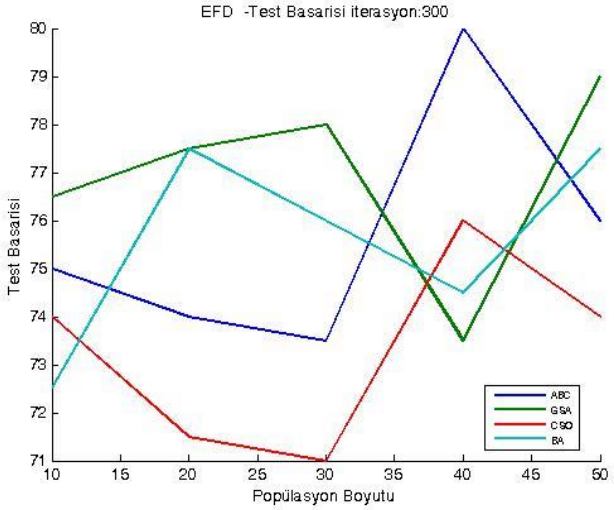
Şekil 3.3. WPBC veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



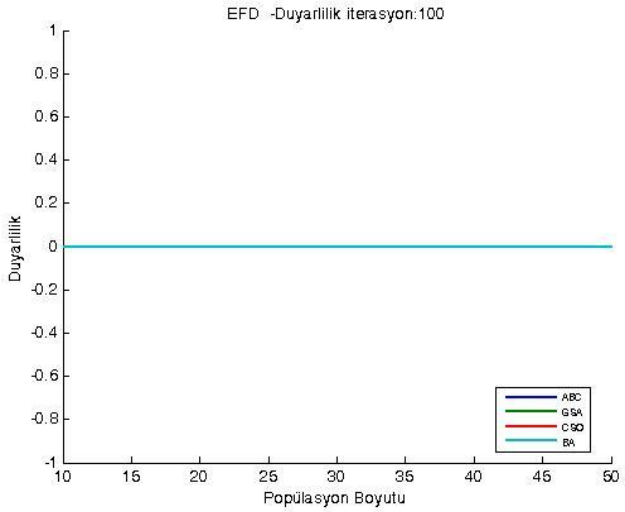
Şekil 3.4. WPBC veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



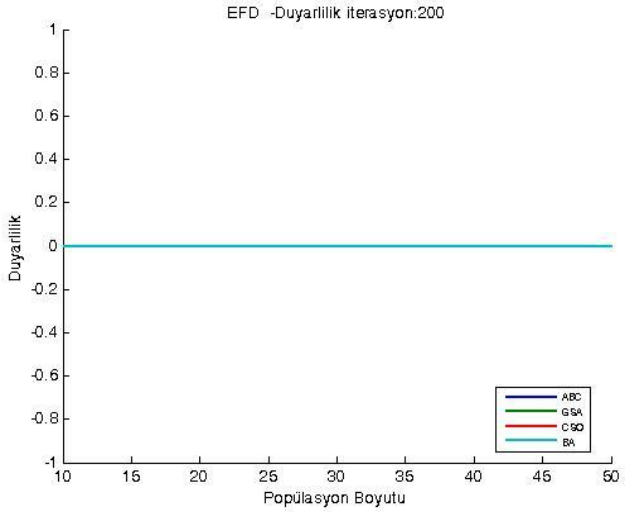
Şekil 3.5. WPBC veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



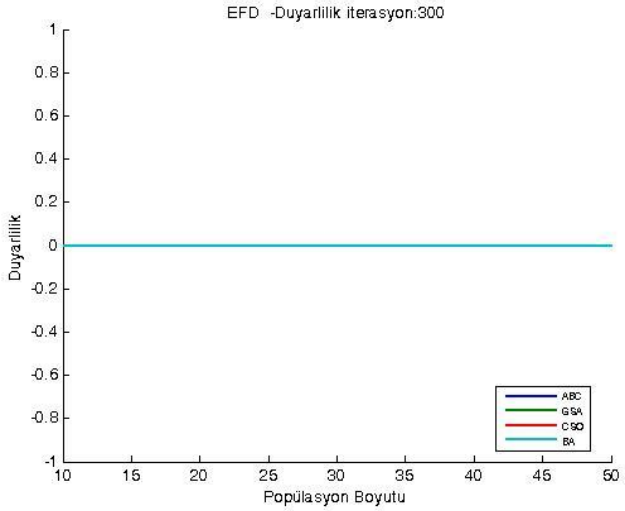
Şekil 3.6. WPBC veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



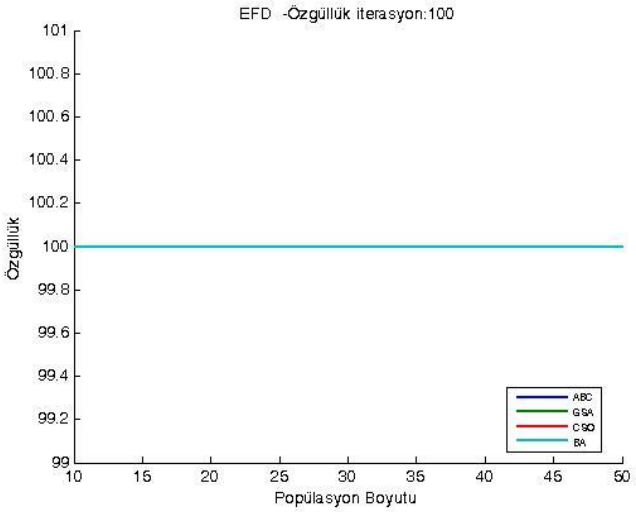
Şekil 3.7. WPBC veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



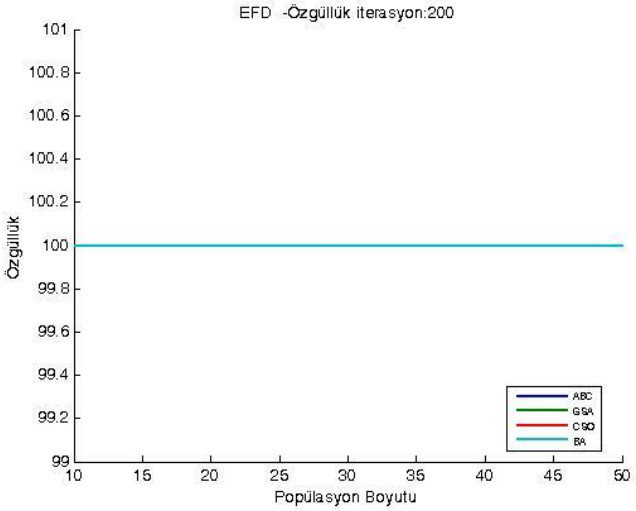
Şekil 3.8. WPBC veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



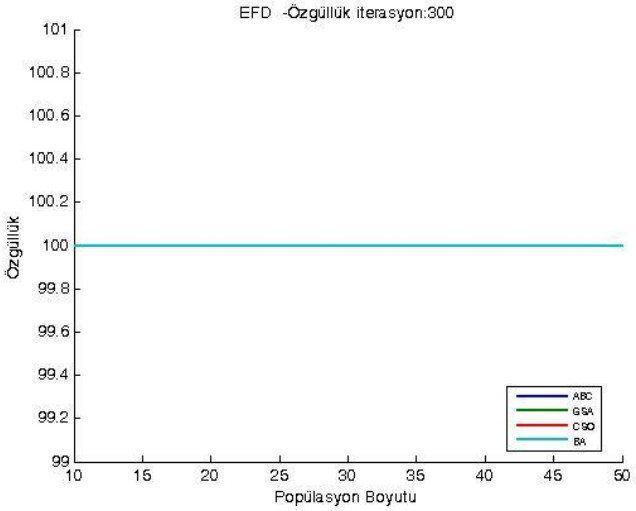
Şekil 3.9. WPBC veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



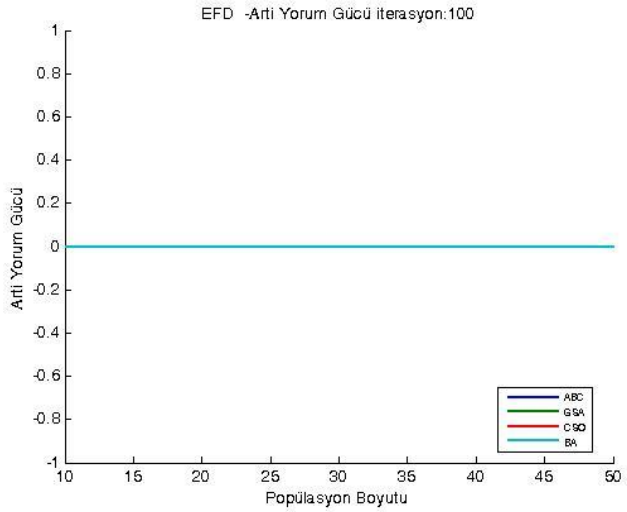
Şekil 3.10. WPBC veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



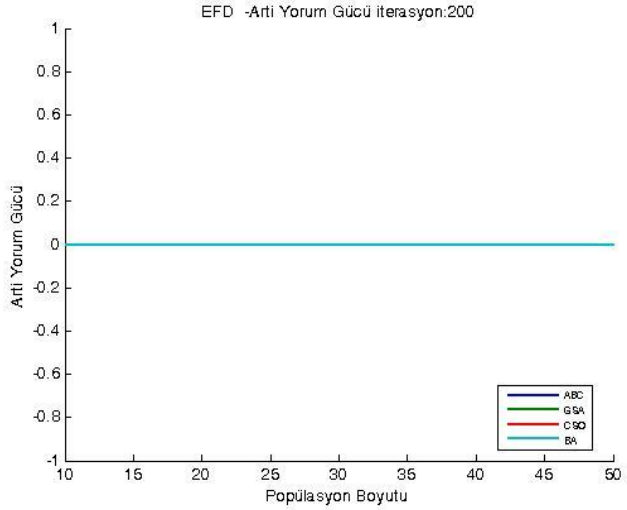
Şekil 3.11. WPBC veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



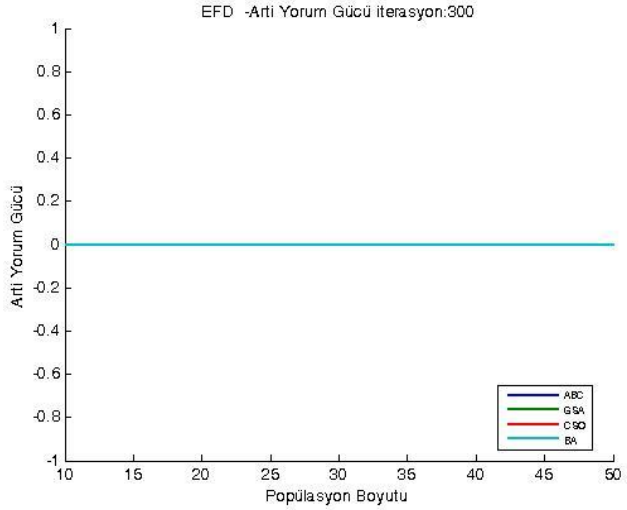
Şekil 3.12. WPBC veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



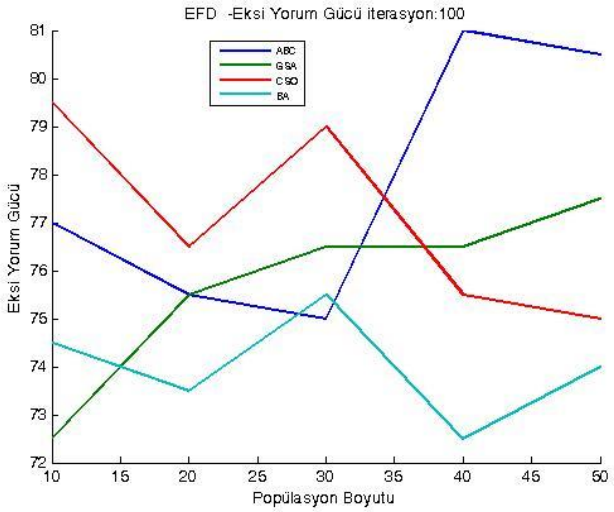
Şekil 3.13. WPBC veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



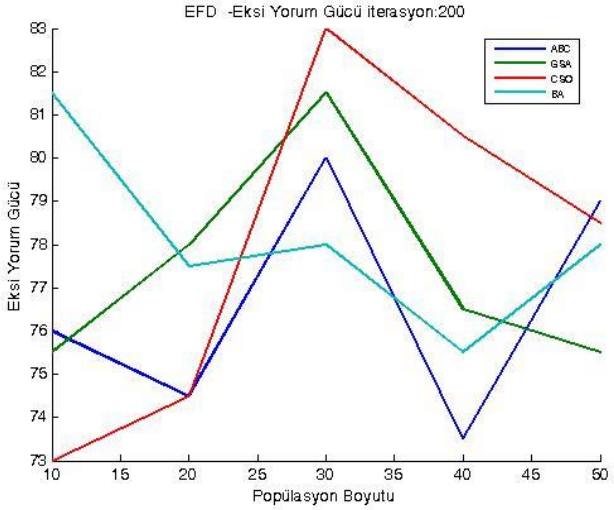
Şekil 3.14. WPBC veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



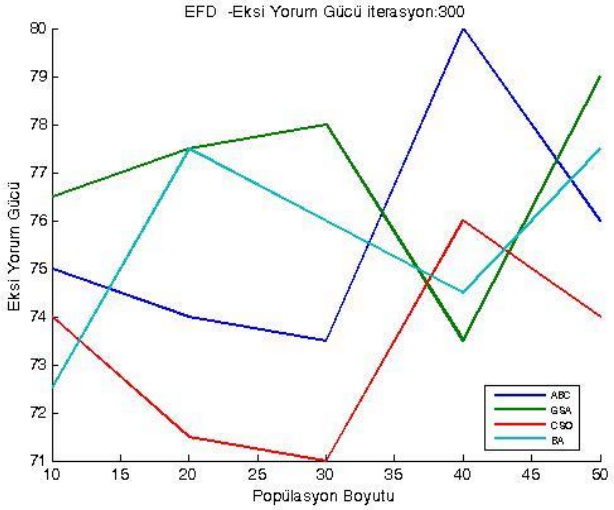
Şekil 3.15. WPBC veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



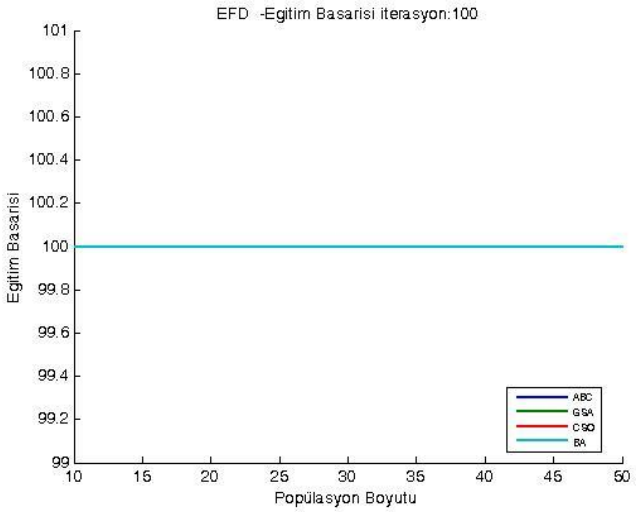
Şekil 3.16. WPBC veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



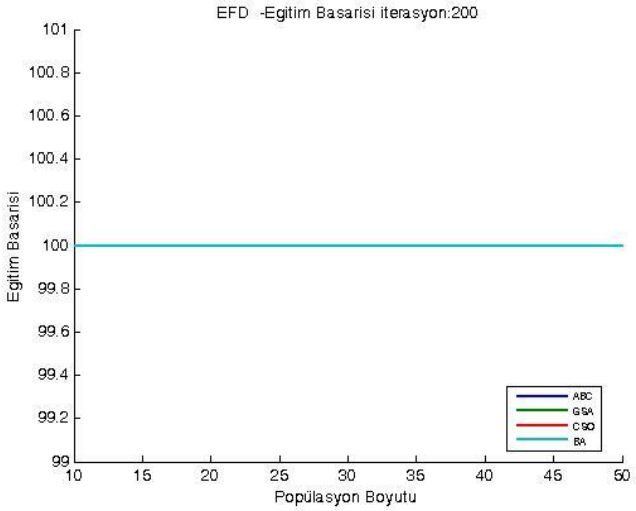
Şekil 3.17. WPBC veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



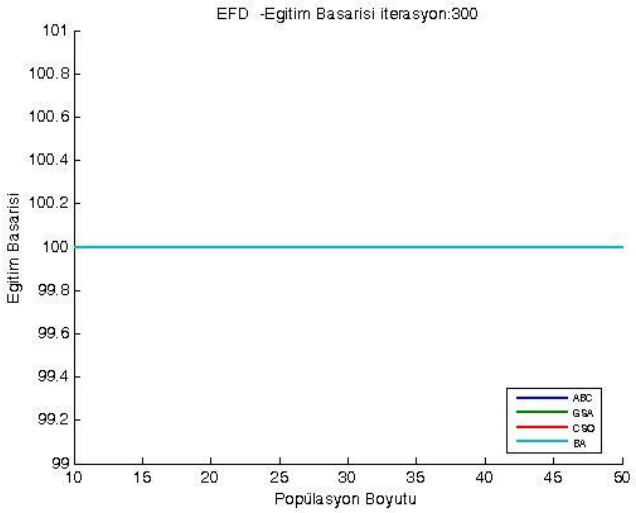
Şekil 3.18. WPBC veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



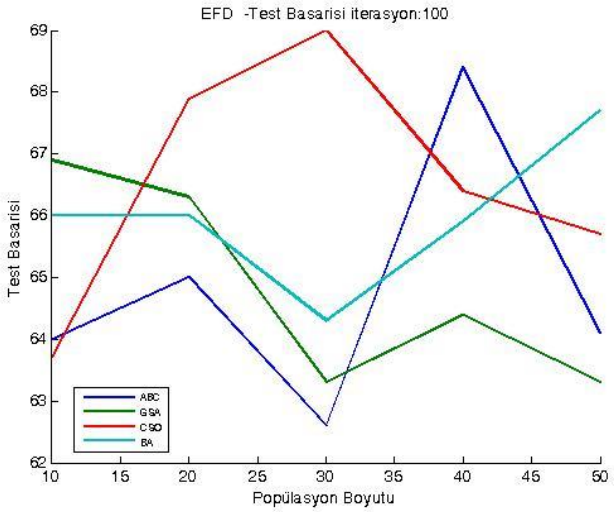
Şekil 3.19. BCW veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



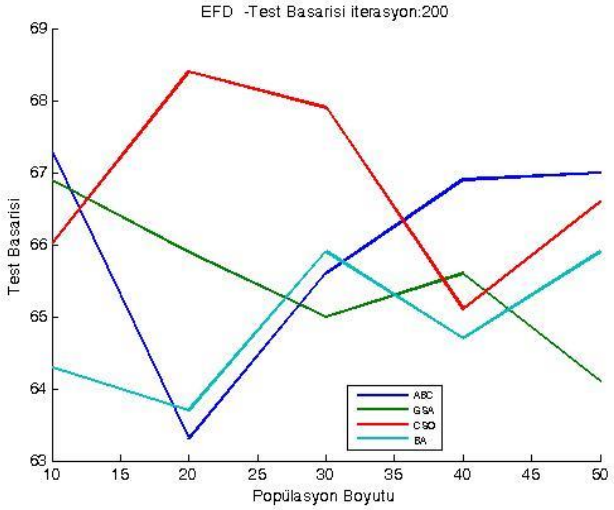
Şekil 3.20. BCW veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



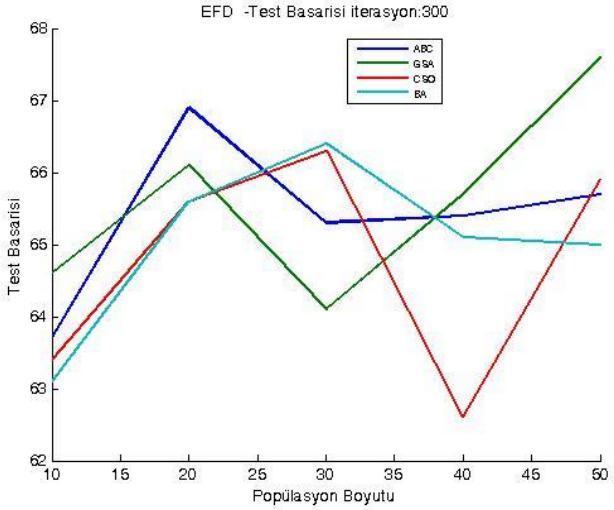
Şekil 3.21. BCW veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



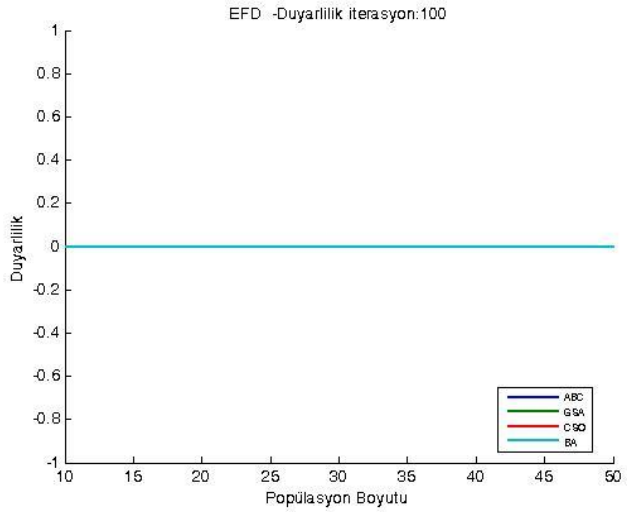
Şekil 3.22. BCW veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



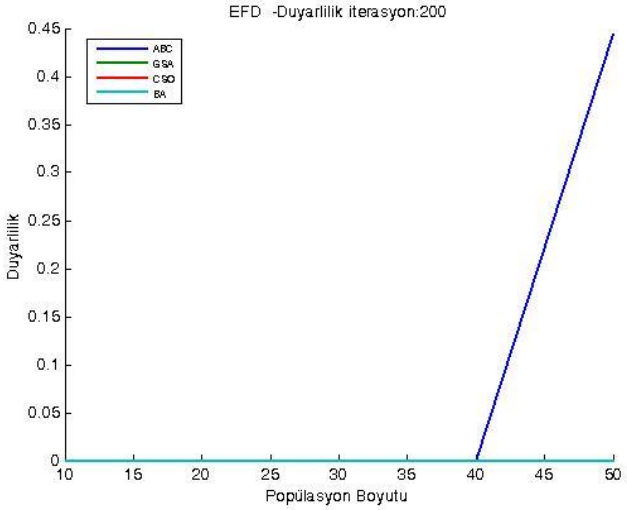
Şekil 3.23. BCW veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



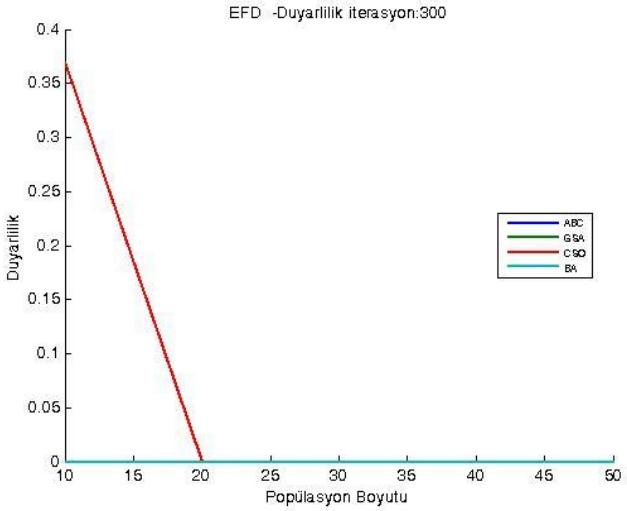
Şekil 3.24. BCW veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



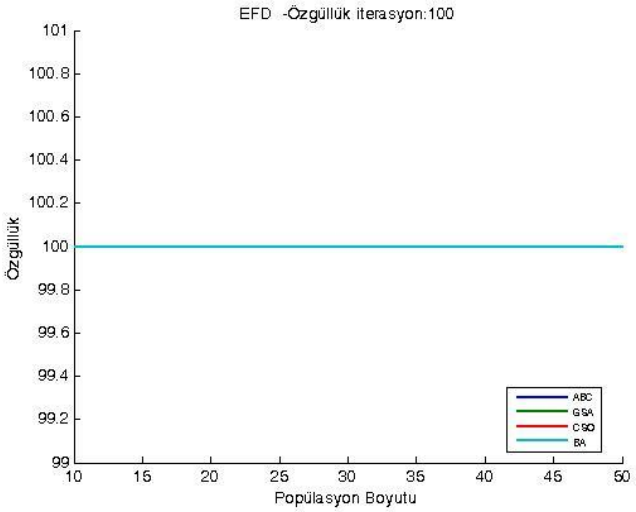
Şekil 3.25. BCW veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



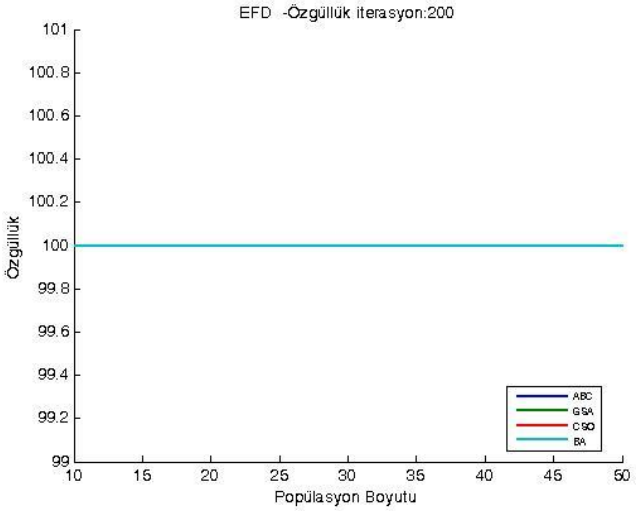
Şekil 3.26. BCW veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



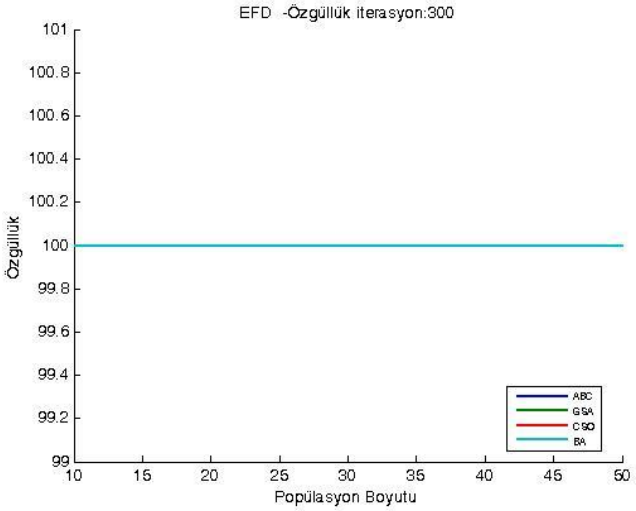
Şekil 3.27. BCW veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



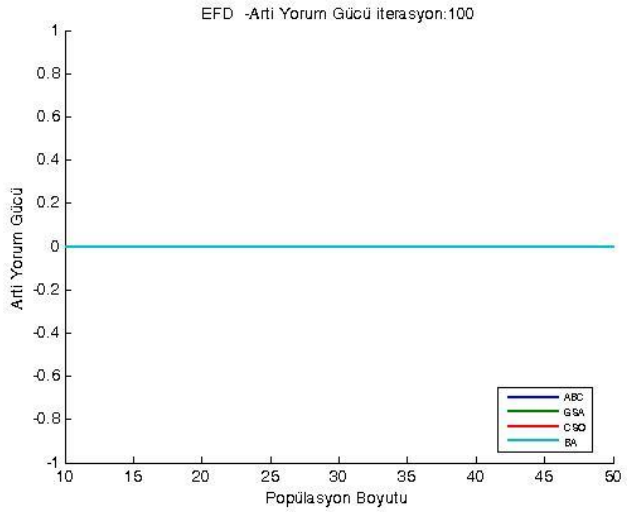
Şekil 3.28. BCW veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



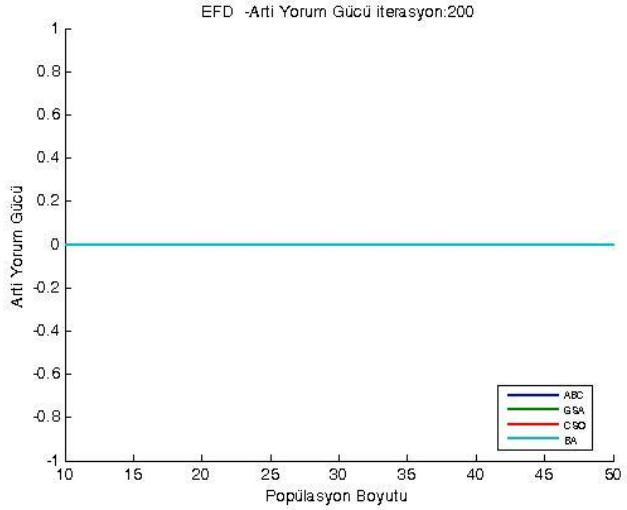
Şekil 3.29. BCW veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



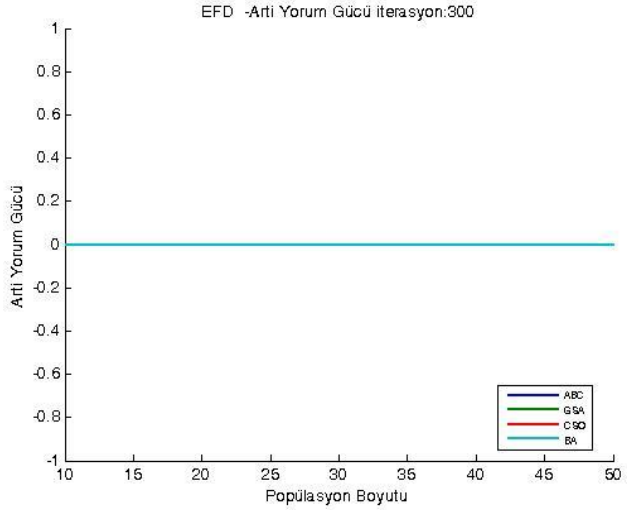
Şekil 3.30. BCW veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



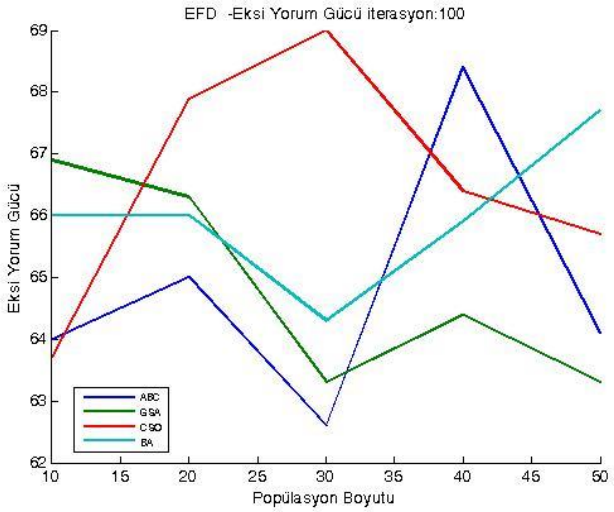
Şekil 3.31. BCW veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



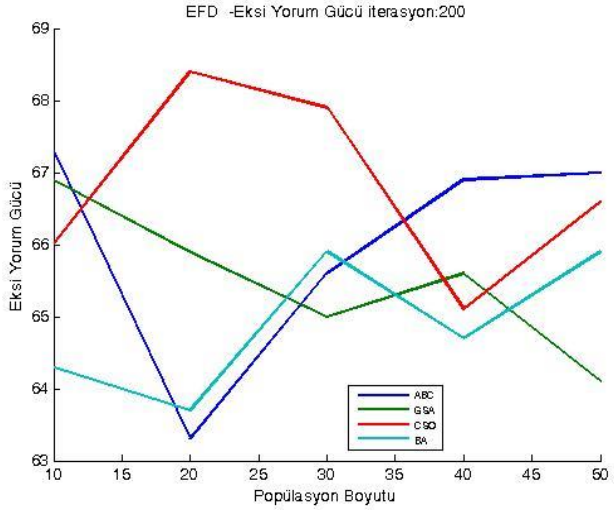
Şekil 3.32. BCW veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



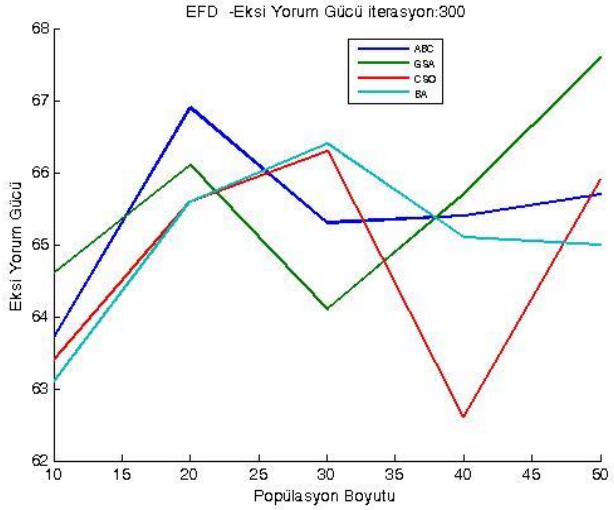
Şekil 3.33. BCW veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



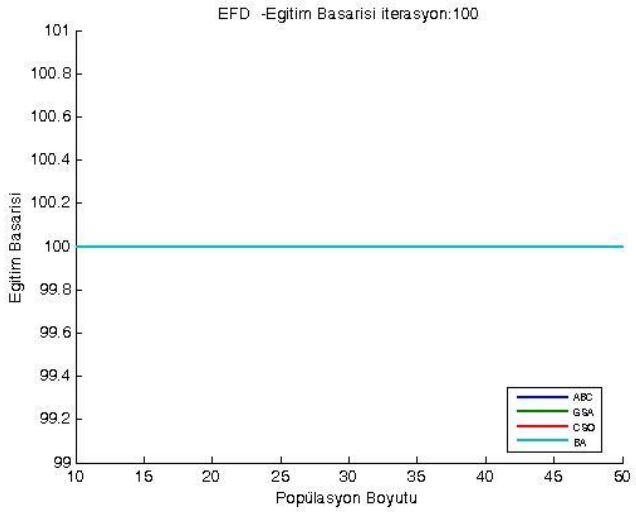
Şekil 3.34. BCW veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



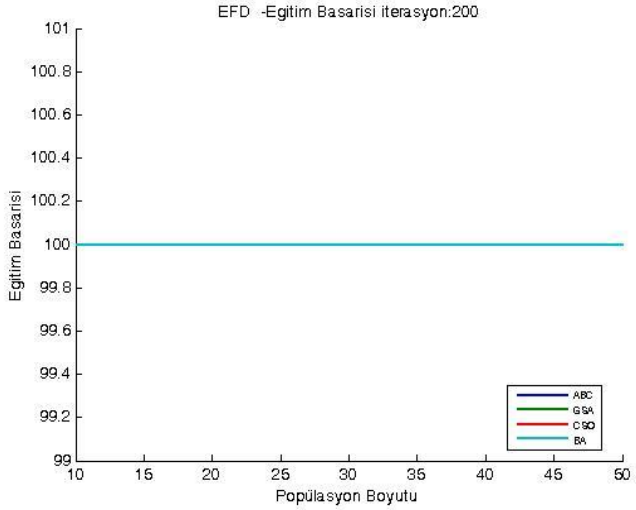
Şekil 3.35. BCW veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



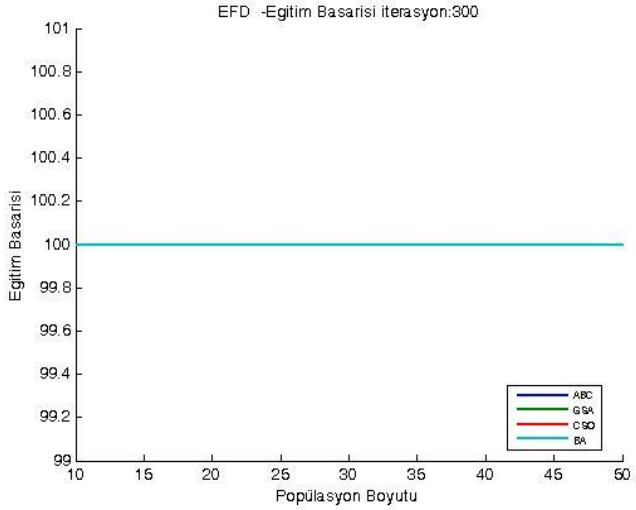
Şekil 3.36. BCW veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



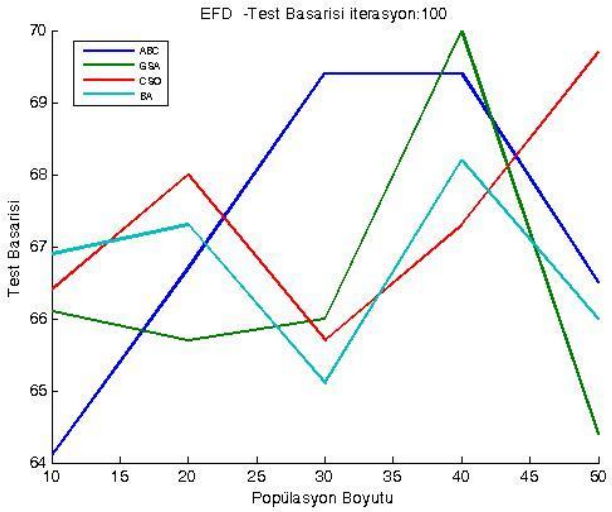
Şekil 3.37. QSAR veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



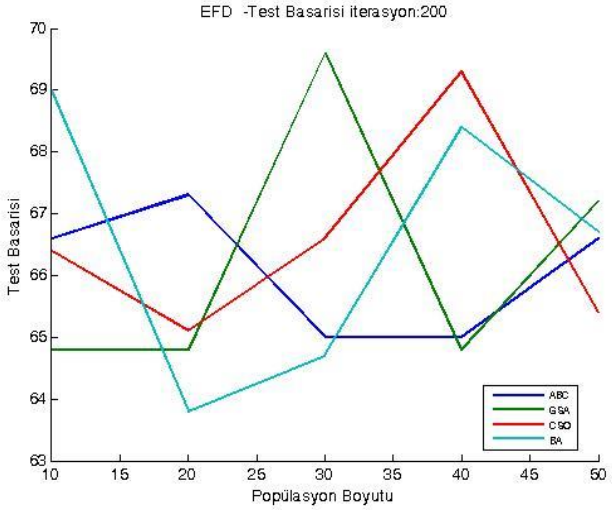
Şekil 3.38. QSAR veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



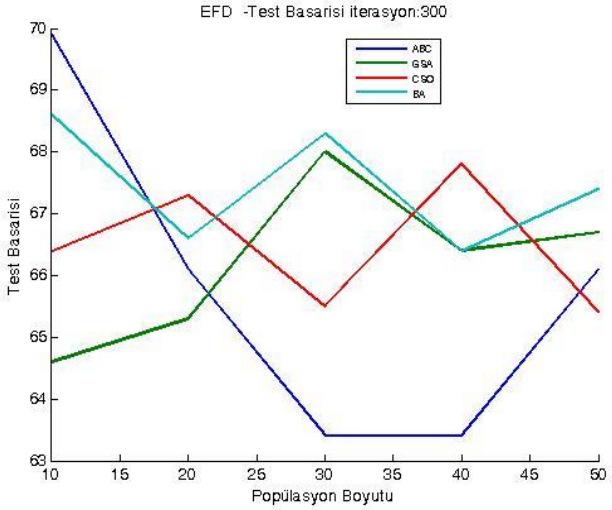
Şekil 3.39. QSAR veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



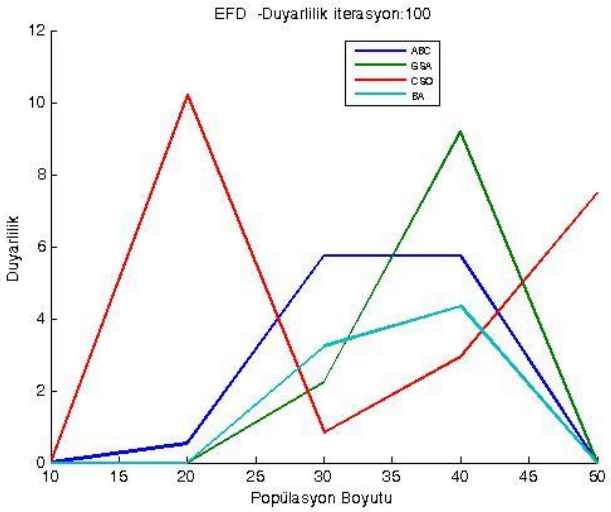
Şekil 3.40. QSAR veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



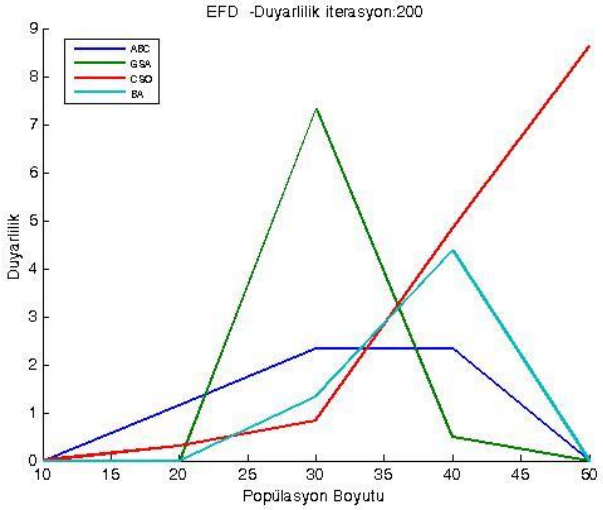
Şekil 3.41. QSAR veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



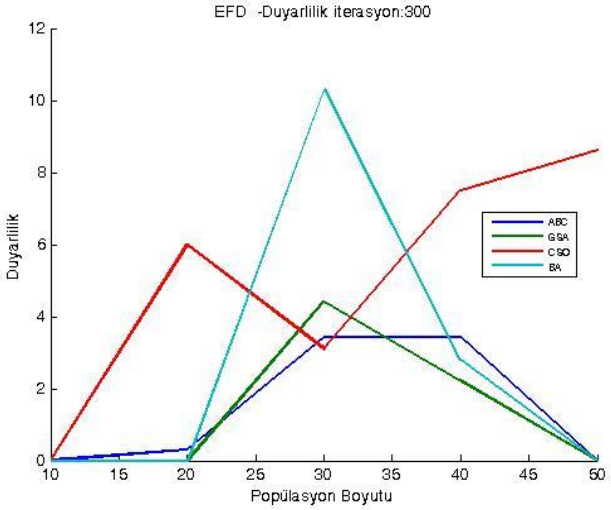
Şekil 3.42. QSAR veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



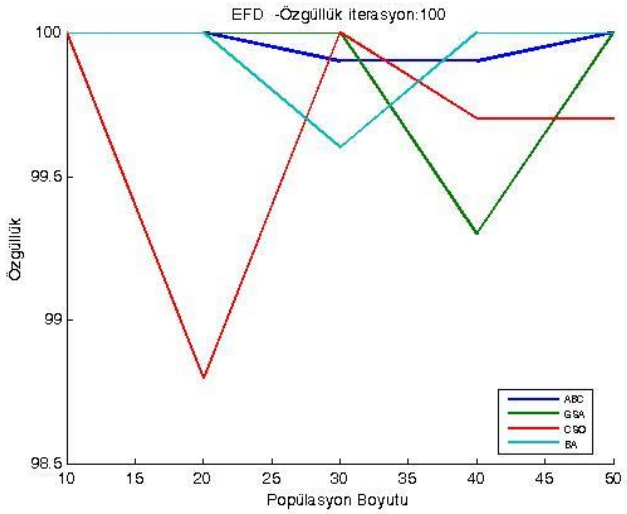
Şekil 3.43. QSAR veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



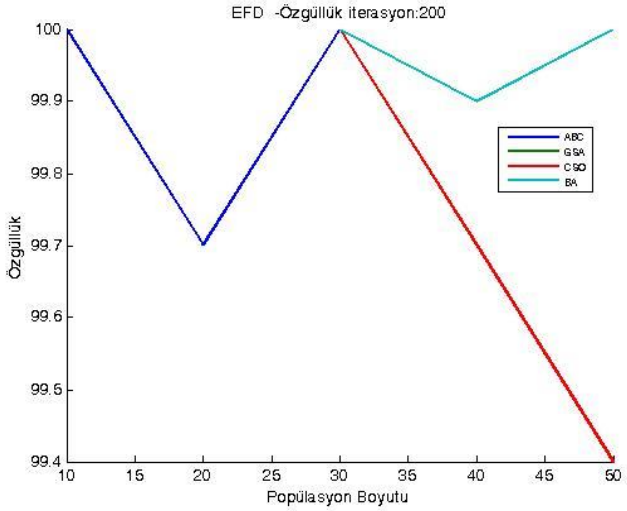
Şekil 3.44. QSAR veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



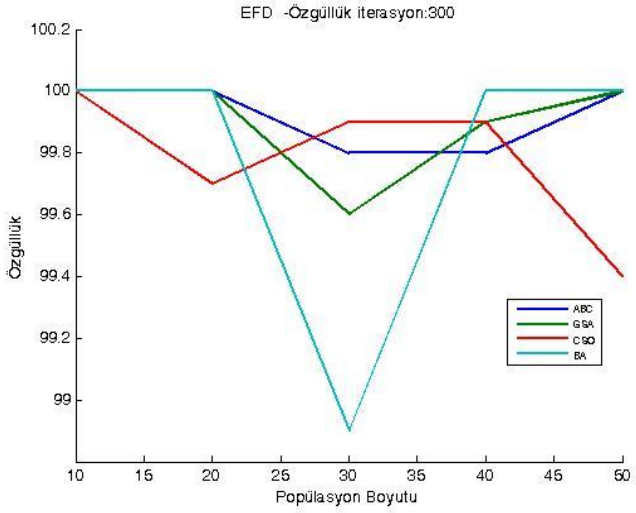
Şekil 3.45. QSAR veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



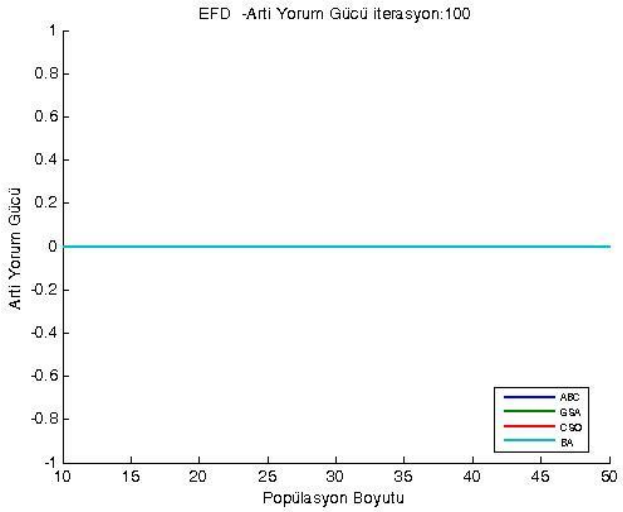
Şekil 3.46. QSAR veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



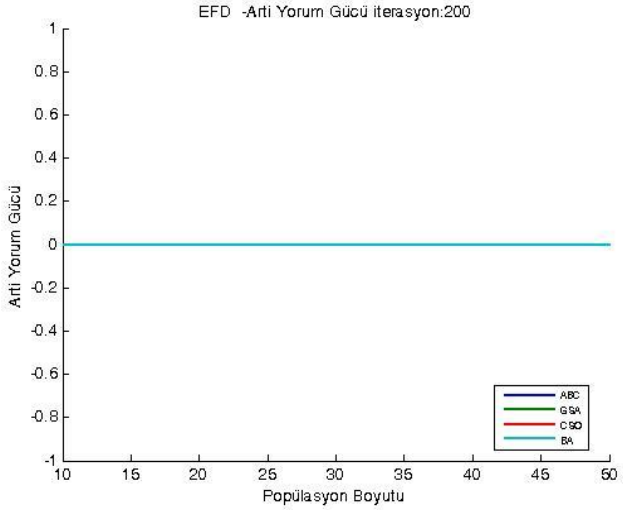
Şekil 3.47. QSAR veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



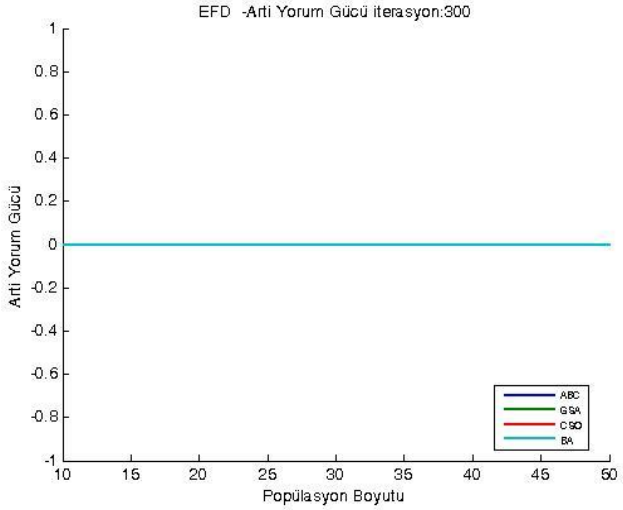
Şekil 3.48. QSAR veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



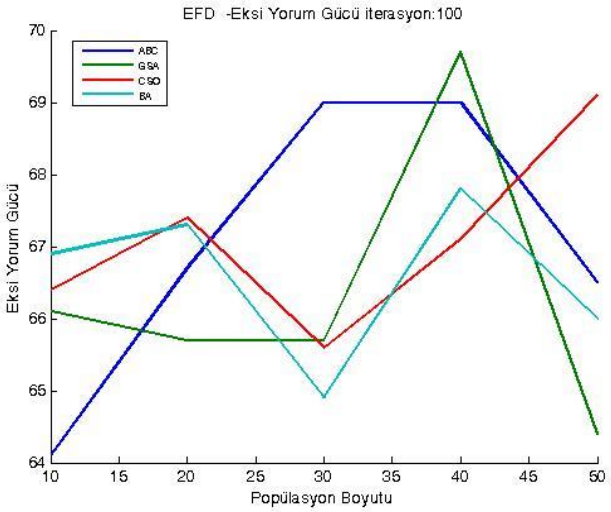
Şekil 3.49. QSAR veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



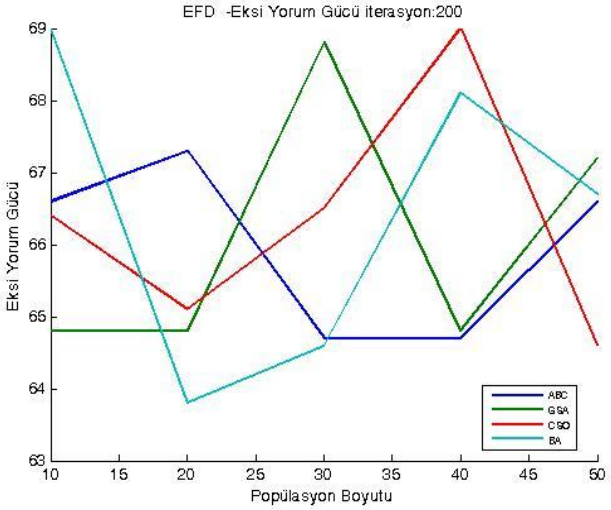
Şekil 3.50. QSAR veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



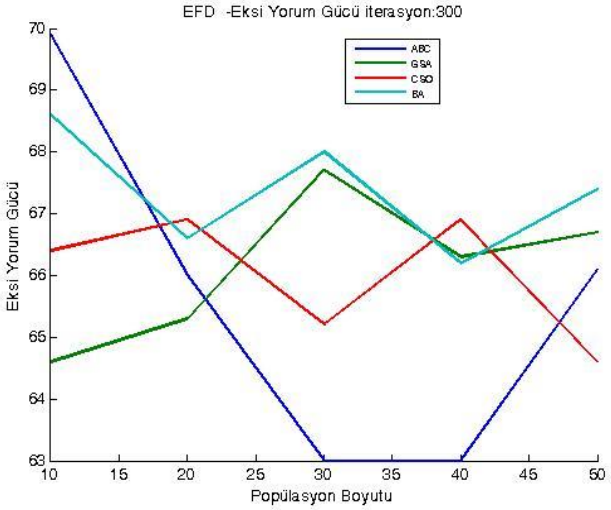
Şekil 3.51. QSAR veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



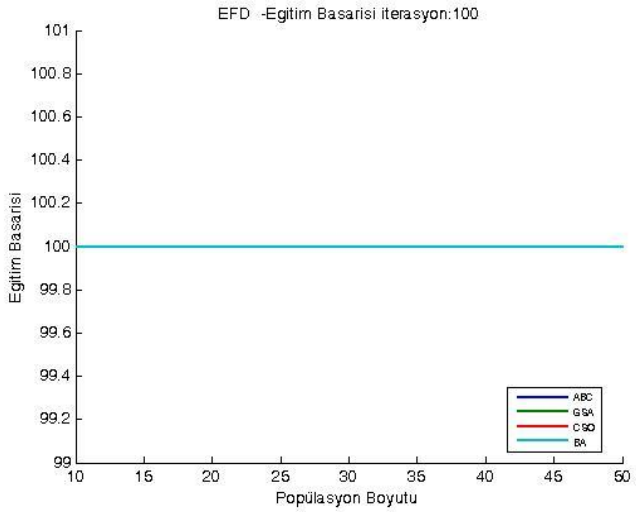
Şekil 3.52. QSAR veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



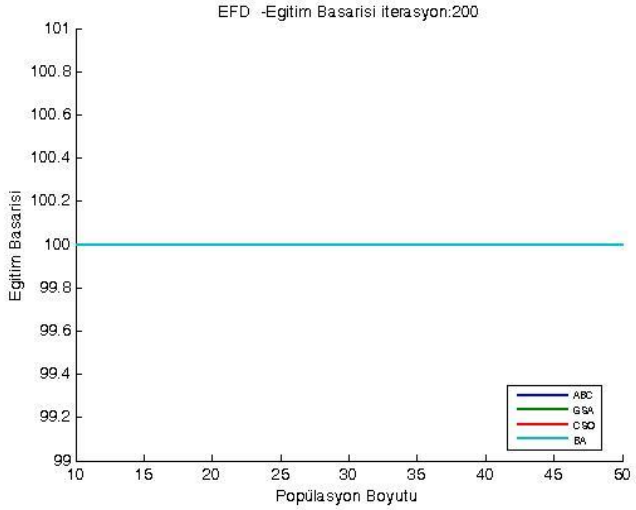
Şekil 3.53. QSAR veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



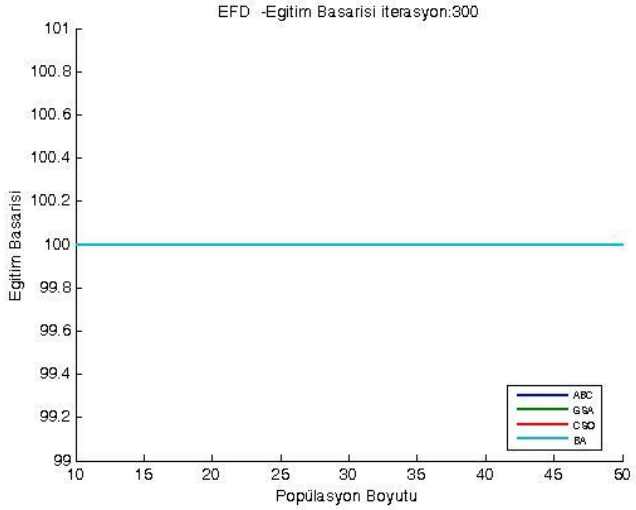
Şekil 3.54. QSAR veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları



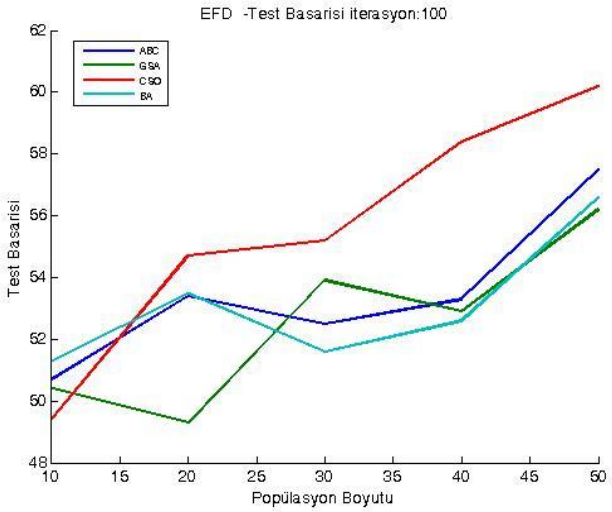
Şekil 3.55. DRD veri kümesi üzerinde 100 iterasyon için ortalama eğitim başarıları



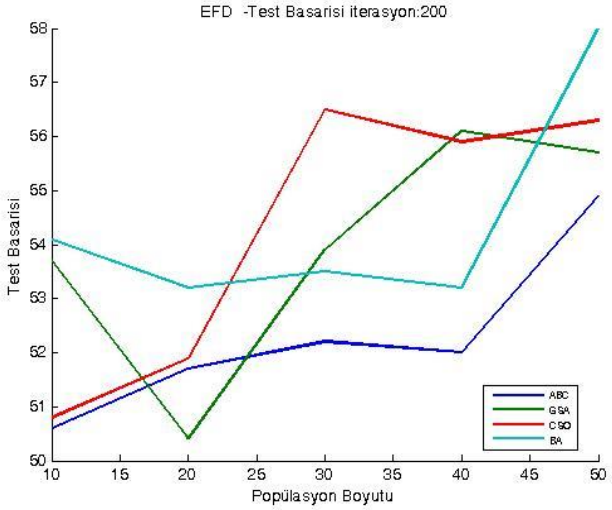
Şekil 3.56. DRD veri kümesi üzerinde 200 iterasyon için ortalama eğitim başarıları



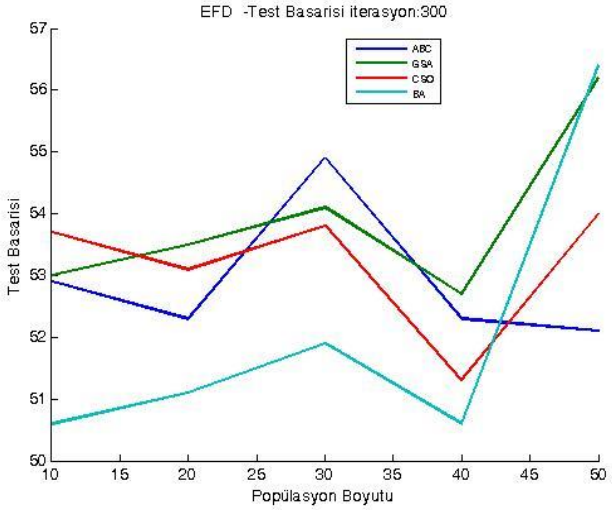
Şekil 3.57. DRD veri kümesi üzerinde 300 iterasyon için ortalama eğitim başarıları



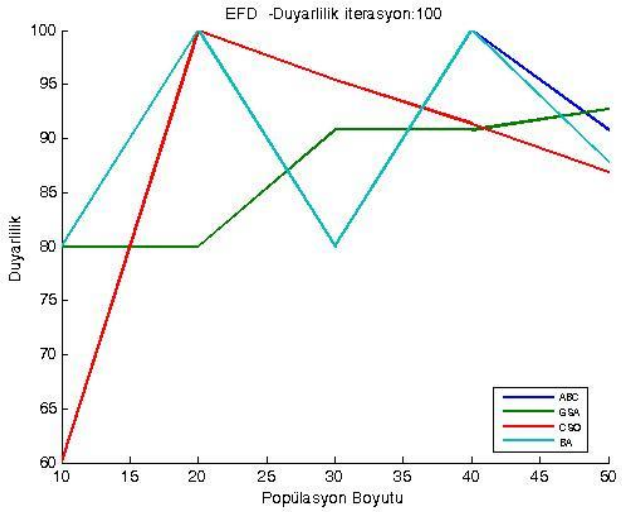
Şekil 3.58. DRD veri kümesi üzerinde 100 iterasyon için ortalama test başarıları



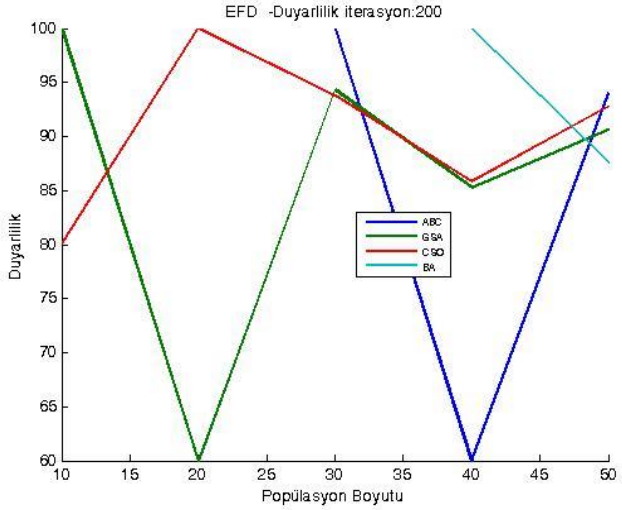
Şekil 3.59. DRD veri kümesi üzerinde 200 iterasyon için ortalama test başarıları



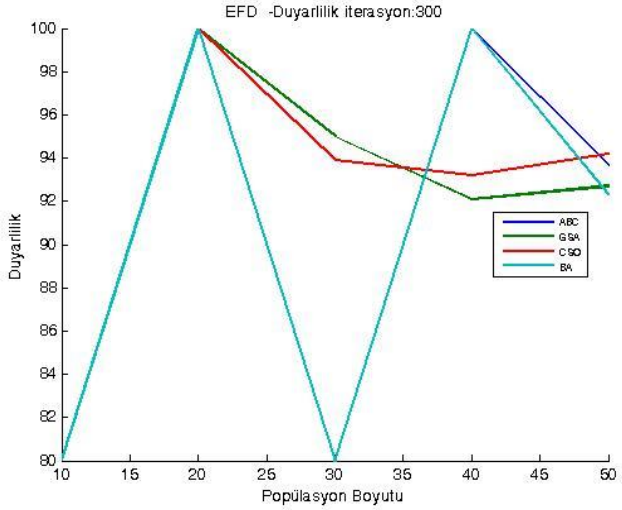
Şekil 3.60. DRD veri kümesi üzerinde 300 iterasyon için ortalama test başarıları



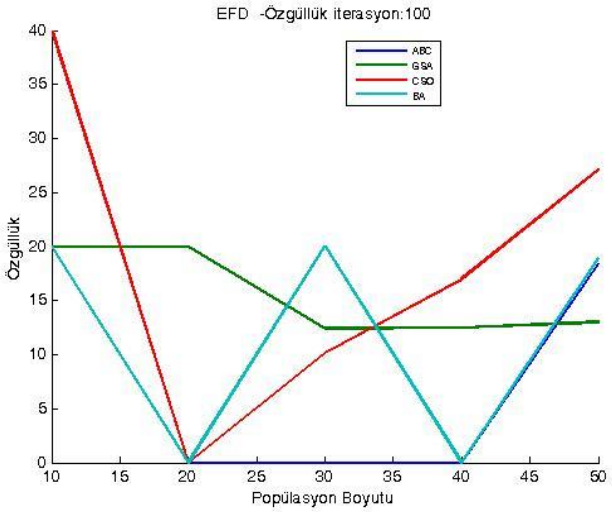
Şekil 3.61. DRD veri kümesi üzerinde 100 iterasyon için ortalama duyarlılık sonuçları



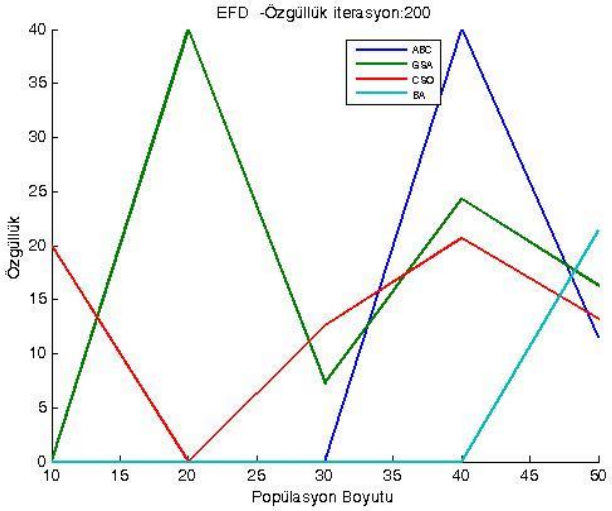
Şekil 3.62. DRD veri kümesi üzerinde 200 iterasyon için ortalama duyarlılık sonuçları



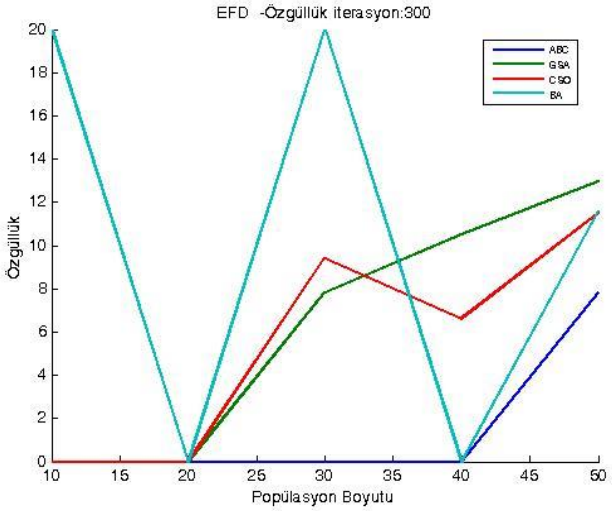
Şekil 3.63. DRD veri kümesi üzerinde 300 iterasyon için ortalama duyarlılık sonuçları



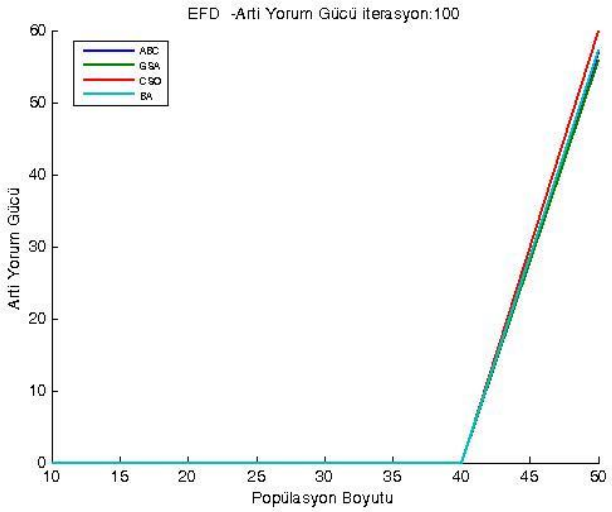
Şekil 3.64. DRD veri kümesi üzerinde 100 iterasyon için ortalama özgüllük sonuçları



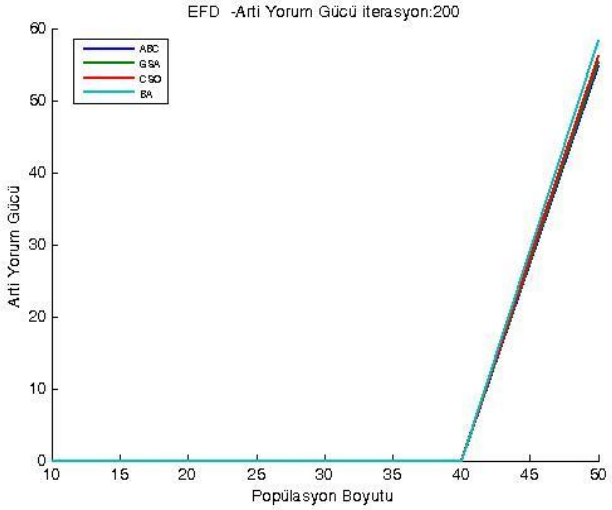
Şekil 3.65. DRD veri kümesi üzerinde 200 iterasyon için ortalama özgüllük sonuçları



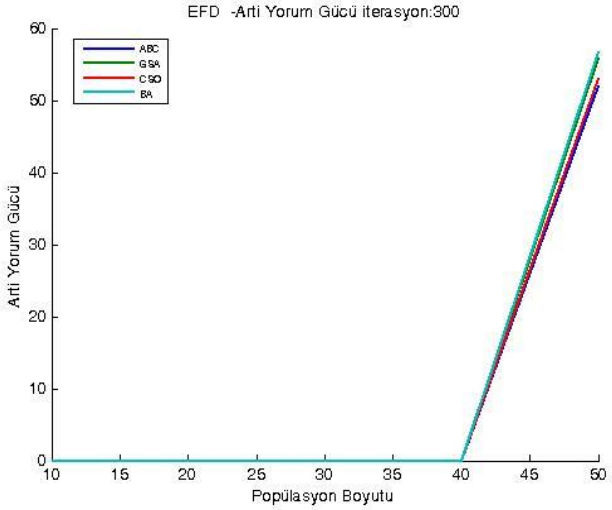
Şekil 3.66. DRD veri kümesi üzerinde 300 iterasyon için ortalama özgüllük sonuçları



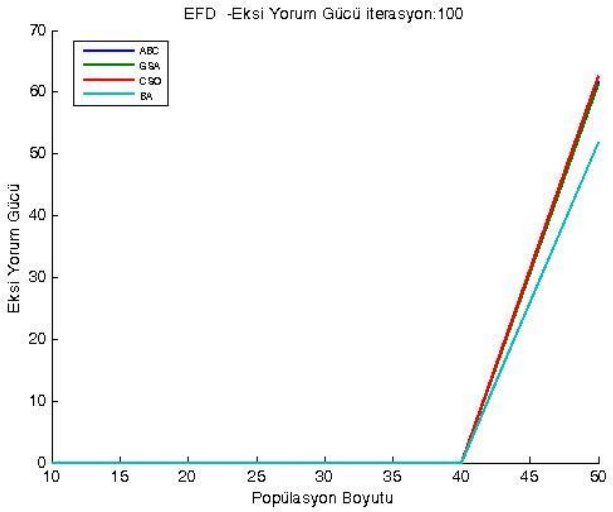
Şekil 3.67. DRD veri kümesi üzerinde 100 iterasyon için ortalama artı yorum gücü sonuçları



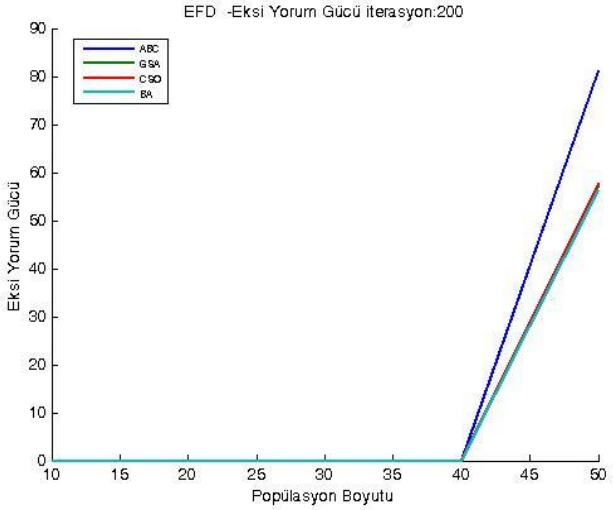
Şekil 3.68. DRD veri kümesi üzerinde 200 iterasyon için ortalama artı yorum gücü sonuçları



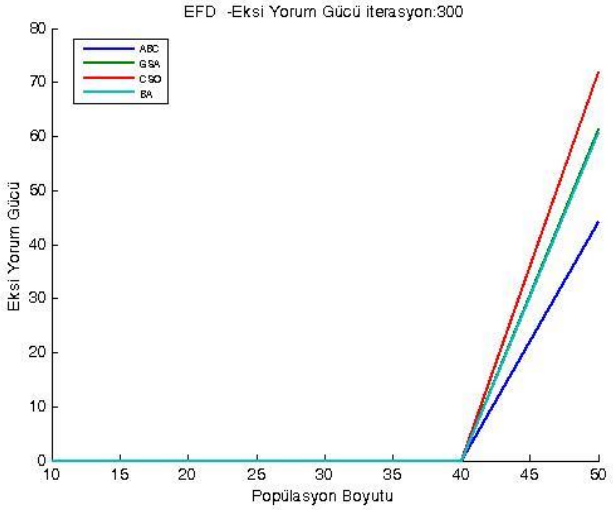
Şekil 3.69. DRD veri kümesi üzerinde 300 iterasyon için ortalama artı yorum gücü sonuçları



Şekil 3.70. DRD veri kümesi üzerinde 100 iterasyon için ortalama eksi yorum gücü sonuçları



Şekil 3.71. DRD veri kümesi üzerinde 200 iterasyon için ortalama eksi yorum gücü sonuçları



Şekil 3.72. DRD veri kümesi üzerinde 300 iterasyon için ortalama eksi yorum gücü sonuçları

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı	**:**	İsmail KOÇ
Uyruğu	**:**	T.C.
Doğum Yeri ve Tarihi	**:**	KONYA 15/02/1987
Telefon	**:**	0 543 648 94 01
Faks	**:**	
e-mail	**:**	ismailkoc@selcuk.edu.tr

EĞİTİM

Derece		**Adı, İlçe, İl**	**Bitirme Yılı**
Lise	**:**	Dumlupınar Lisesi, Selçuklu, Konya	2004
Üniversite	**:**	Selçuk Üniversitesi, Selçuklu, Konya	2010
Yüksek Lisans:		Selçuk Üniversitesi, Selçuklu, Konya	-

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**
2011	Elpo Bilişim Otomasyon	Bilgisayar Mühendisi
2012	Ayosis Yazılım	Bilgisayar Mühendisi
2013	Hitit Üniversitesi	Araştırma Görevlisi

UZMANLIK ALANI

Sınıflandırma, Yapay Zeka, Makina Öğrenmesi, Optimizasyon Yöntemleri

YABANCI DİLLER

İngilizce

YAYINLAR

Koc I, Babaoglu I, 2015, A novel modified binary bat algorithm for feature selection in classification problems, International Conference on Advanced Technology & Sciences, 2th International Conference, ICAT'15 Antalya, Turkey, August 04-07, 2015, 195-200.